



Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Estudios de Postgrado
Maestría en Estadística Aplicada

**MODELO DE REGRESIÓN MULTIVARIADO APLICADO A LAS CONDICIONES DE
DESNUTRICIÓN INFANTIL, PARA MEJORAR LAS PROPUESTAS DE DESARROLLO
SOCIAL EN EL SUROCCIDENTE DE GUATEMALA**

Ing. Daniel Josué Granados Mota

Asesorado por el M.A. Adonai Navas García

Guatemala, septiembre de 2023

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**MODELO DE REGRESIÓN MULTIVARIADO APLICADO A LAS CONDICIONES DE
DESNUTRICIÓN INFANTIL, PARA MEJORAR LAS PROPUESTAS DE DESARROLLO
SOCIAL EN EL SUROCCIDENTE DE GUATEMALA**

TRABAJO DE GRADUACIÓN

PRESENTADO A LA JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERÍA
POR

ING. DANIEL JOSUÉ GRANADOS MOTA
ASESORADO POR EL M.A. ADONAI NAVAS GARCÍA

AL CONFERÍRSELE EL TÍTULO DE

MAESTRO EN ESTADÍSTICA APLICADA

GUATEMALA, SEPTIEMBRE DE 2023

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA



NÓMINA DE JUNTA DIRECTIVA

DECANO	Ing. José Francisco Gómez Rivera (a.i)
VOCAL II	Ing. Mario Renato Escobedo Martínez
VOCAL III	Ing. José Milton de León Bran
VOCAL IV	Ing. Kevin Vladimir Armando Cruz
VOCAL V	Br. Fernando José Paz González
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

**TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL DE DEFENSA DE
TESIS**

DECANA	Inga. Aurelia Anabela Cordova Estrada
EXAMINADOR	Inga. Aurelia Anabela Cordova Estrada
EXAMINADOR	Ing. Edwin Adalberto Bracamonte Orozco
EXAMINADOR	Ing. William Eduardo Fagiani Cruz
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

HONORABLE TRIBUNAL EXAMINADOR

En cumplimiento con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

**MODELO DE REGRESIÓN MULTIVARIADO APLICADO A LAS CONDICIONES DE
DESNUTRICIÓN INFANTIL, PARA MEJORAR LAS PROPUESTAS DE DESARROLLO
SOCIAL EN EL SUROCCIDENTE DE GUATEMALA**

Tema que me fuera asignado por la Dirección de la Escuela de Estudios de Postgrado, con fecha 31 de enero de 2022.



Ing. Daniel Josué Granados Mota

LNG.DECANATO.OI.625.2023

El Decano de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Estudios de Posgrado, al Trabajo de Graduación titulado: **MODELO DE REGRESIÓN MULTIVARIADO APLICADO A LAS CONDICIONES DE DESNUTRICIÓN INFANTIL, PARA MEJORAR LAS PROPUESTAS DE DESARROLLO SOCIAL EN EL SUROCCIDENTE DE GUATEMALA**, presentado por: **Ing. Daniel Josué Granados Mota**, que pertenece al programa de Maestría en artes en Estadística aplicada después de haber culminado las revisiones previas bajo la responsabilidad de las instancias correspondientes, autoriza la impresión del mismo.

IMPRÍMASE:



Ing. José Francisco Gómez Rivera

Decano a.i.

Guatemala, septiembre de 2023

JFGR/gaoc



Guatemala, septiembre de 2023

LNG.EEP.OI.625.2023

En mi calidad de Directora de la Escuela de Estudios de Postgrado de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer el dictamen del asesor, verificar la aprobación del Coordinador de Maestría y la aprobación del Área de Lingüística al trabajo de graduación titulado:

“MODELO DE REGRESIÓN MULTIVARIADO APLICADO A LAS CONDICIONES DE DESNUTRICIÓN INFANTIL, PARA MEJORAR LAS PROPUESTAS DE DESARROLLO SOCIAL EN EL SUROCCIDENTE DE GUATEMALA”

presentado por **Ing. Daniel Josué Granados Mota** correspondiente al programa de **Maestría en artes en Estadística aplicada** ; apruebo y autorizo el mismo.

Atentamente,

“Id y Enseñad a Todos”



Mtra. Inga. Aurelia Anabela Cordova Estrada
Directora
Escuela de Estudios de Postgrado
Facultad de Ingeniería



Guatemala 30 de mayo 2022.

M.A. Edgar Darío Álvarez Cotí
Director
Escuela de Estudios de Postgrado
Presente

M.A. Ingeniero Álvarez Cotí:

Por este medio informo que he revisado y aprobado el **INFORME FINAL y ARTICULO CIENTÍFICO** titulado: **MODELO DE REGRESIÓN MULTIVARIADO APLICADO A LAS CONDICIONES DE DESNUTRICIÓN INFANTIL, PARA MEJORAR LAS PROPUESTAS DE DESARROLLO SOCIAL EN EL SUROCCIDENTE DE GUATEMALA**, del estudiante **Daniel Josué Granados Mota** quien se identifica con número de carné **201130100** del programa de Maestría en Estadística Aplicada.

Con base en la evaluación realizada hago constar que he evaluado la calidad, validez, pertinencia y coherencia de los resultados obtenidos en el trabajo presentado y según lo establecido en el *Normativo de Tesis y Trabajos de Graduación aprobado por Junta Directiva de la Facultad de Ingeniería Punto Sexto inciso 6.10 del Acta 04-2014 de sesión celebrada el 04 de febrero de 2014*. Por lo cual el trabajo evaluado cuenta con mi aprobación.

Agradeciendo su atención y deseándole éxitos en sus actividades profesionales me suscribo.

Atentamente,




MSc. Ing. Edwin Adalberto Bracamonte Orozco
Coordinador
Maestría Estadística Aplicada
Escuela de Estudios de Postgrado

Guatemala, mayo de 2022.

Mtro. Ing. Edgar Darío Álvarez Cotí.
Director de la Escuela de Estudios de Postgrado. FIUSAC.
Presente.

Estimado Maestro Álvarez Cotí:

De la manera más atenta me dirijo a usted, esperando que sus actividades se desarrollen con éxito, y a la vez, hago de su conocimiento que he revisado y aprobado el Trabajo de Graduación y el Artículo Científico: **“MODELO DE REGRESIÓN MULTIVARIADO APLICADO A LAS CONDICIONES DE DESNUTRICIÓN INFANTIL, PARA MEJORAR LAS PROPUESTAS DE DESARROLLO SOCIAL EN EL SUROCCIDENTE DE GUATEMALA”**, del estudiante **Daniel Josué Granados Mota** del programa de Maestría en **Estadística Aplicada**, identificado con el carné: **2520419250901**

Sin otro particular, me suscribo a sus respetables órdenes.

Atentamente,



M.A. Lic. Adonai Navas García
Asesor de Tesis

ACTO QUE DEDICO A:

Dios

Por ser una importante influencia en mi carrera, entre otras cosas.

Mis padres

Carlos Granados y Mercedes Mota.
Su amor será siempre mi inspiración.

AGRADECIMIENTOS A:

**Universidad de San Carlos de
Guatemala**

Por ser una importante influencia en
mi carrera, entre otras cosas.

Facultad de Ingeniería

Por ser una importante influencia en
mi carrera, entre otras cosas.

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES.....	VII
LISTA DE SÍMBOLOS	XI
GLOSARIO	XIII
RESUMEN.....	XVII
PLANTEAMIENTO DEL PROBLEMA.....	XIX
OBJETIVOS.....	XXIII
RESUMEN DEL MARCO METODOLÓGICO	XXV
INTRODUCCIÓN	XXXVII
1. MARCO REFERENCIAL.....	1
2. MARCO TEÓRICO.....	9
2.1. Fundamentos estadísticos.....	9
2.1.1. Regresión múltiple	9
2.1.2. Supuestos estadísticos de la regresión múltiple	10
2.1.2.1. Independencia	11
2.1.2.2. Homocedasticidad	12
2.1.2.3. Multicolinealidad	13
2.1.2.4. Prueba de hipótesis	15
2.1.2.5. Principio de parsimonia	20
2.1.3. Regresión logística binomial	21
2.1.4. Validadores de la regresión logística	22
2.1.4.1. Análisis de devianzas	22
2.1.4.2. Estadístico de chi cuadrado.....	23
2.1.4.3. Estadístico de Wald.....	24

	2.1.4.4.	Odds Ratio.....	26
	2.1.4.5.	Criterio de información AIC.....	28
	2.1.4.6.	Prueba de Hosmer Lemeshow	29
	2.1.5.	Plan de muestreo.....	31
2.2.		Indicadores nutricionales infantiles y variables RV	32
	2.2.1.	Medición nutricional infantil.....	32
	2.2.1.1.	Malnutrición infantil	33
	2.2.1.2.	Clasificación del estado nutricional infantil	34
	2.2.1.3.	Utilización de índices antropométricos...36	
	2.2.2.	Características de vivienda e higiene (variables RV)	43
	2.2.2.1.	Definición y morfología de vivienda	44
	2.2.2.2.	Sistemas y métodos constructivos.....	44
	2.2.2.3.	Abastecimiento de agua salubre.....	47
	2.2.2.4.	Eliminación higiénica de excretas.....	48
	2.2.2.5.	Eliminación de desechos sólidos.....	49
	2.2.2.6.	Salvaguardias estructurales.....	50
3.		PRESENTACIÓN DE RESULTADOS.....	53
	3.1.	Objetivo 1. Determinación de la medida en la que se relacionan estadísticamente la presencia de desnutrición infantil con las características físicas y energéticas de los hogares, para mejorar las propuestas de desarrollo social mediante el uso de análisis de devianza.	53
	3.1.1.	Cálculo de diagnóstico.....	53
	3.1.1.1.	Diagnóstico de Waterlow	54
	3.1.1.2.	Diagnóstico según desviaciones estándar de la OMS.....	59

3.1.2.	Análisis de devianzas para la relación entre variables independientes con la presencia de desnutrición infantil	63
3.2.	Objetivo 2. Explicación del comportamiento de las condiciones de desnutrición infantil, en relación con la información correspondiente a los hogares, para mejorar las propuestas de desarrollo social por medio de un modelo de regresión logística multivariada en el suroccidente de Guatemala.....	70
3.2.1.	Desarrollo de modelos de regresión logística para variables respuesta.....	71
3.2.1.1.	Modelo de regresión logística para variable ZTE	72
3.2.1.2.	Modelo de regresión logística para variable ZPT	75
3.2.1.3.	Modelo de regresión logística para variable W.....	79
3.3.	Objetivo 3. Determinación de los factores físicos correspondientes a los hogares que son significativos en las condiciones de desnutrición infantil, y en qué niveles estos factores provocan condiciones de desnutrición, para mejorar las propuestas de desarrollo social mediante un modelo de regresión logística multivariada y el análisis de odds ratios en el suroccidente de Guatemala	82
3.3.1.	Factores y Odds Ratios de niveles significativos para la variable de respuesta ZTE	83
3.3.2.	Factores y Odds Ratios de niveles significativos para la variable de respuesta ZPT	84

3.3.3.	Factores y Odds Ratios de niveles significativos para la variable de respuesta W	84
4.	DISCUSIÓN DE RESULTADOS	88
4.1.	Análisis interno	88
4.1.1.	Discusión de variables significativas relacionadas a infraestructura de vivienda con la presencia de desnutrición infantil	88
4.1.2.	Validación y comparación de modelos de regresión logística desarrollados	89
4.1.2.1.	Validación de modelos según devianza específica	90
4.1.2.2.	Validación de modelos según prueba chi cuadrado	92
4.1.2.3.	Validación de modelos según estadístico de Wald	94
4.1.2.4.	Validación de modelos según criterio de información de Akaike (AIC)	95
4.1.2.5.	Validación de modelos según prueba de Hosmer Lemeshow	96
4.1.3.	Discusión de variables y niveles significativos como causas de condiciones de desnutrición infantil	97
4.1.3.1.	Comparación de variables significativas según probabilidad de presencia de desnutrición	97
4.1.3.2.	Análisis de comportamiento específico de variables de respuesta por factor significativo	98

4.2.	Análisis externo	105
	CONCLUSIONES	109
	RECOMENDACIONES	111
	REFERENCIAS	113
	APÉNDICES	119
	ANEXOS	125

ÍNDICE DE ILUSTRACIONES

FIGURAS

Figura 1.	Clasificación nutricional por Waterlow.....	35
Figura 2.	Longitud para la edad niños.....	38
Figura 3.	Peso para la edad niños	39
Figura 4.	Peso para la talla niños.....	41
Figura 5.	Aumento de probabilidad de desnutrición ZTE según variable fuente_agua.....	99
Figura 6.	Aumento de probabilidad de desnutrición ZTE según variable disposición_excretas.....	100
Figura 7.	Aumento de probabilidad de desnutrición ZPT según variable fuente_agua.....	101
Figura 8.	Aumento de probabilidad de desnutrición W según variable fuente_agua.....	102
Figura 9.	Aumento de probabilidad de desnutrición W según variable piso_casa.....	103
Figura 10.	Aumento de probabilidad de desnutrición W según variable disposición_excretas.....	104

TABLAS

Tabla 1.	Definición de variables.....	XXVI
Tabla 2.	Niveles de variable Material piso	XXVIII
Tabla 3.	Niveles de variable Tipo casa	XXVIII

Tabla 4.	Niveles de variable Fuente agua	XXIX
Tabla 5.	Niveles de variable Calidad agua	XXIX
Tabla 6.	Niveles de variable Combustible cocina	XXX
Tabla 7.	Niveles de variable Disposición basura	XXX
Tabla 8.	Niveles de variable Disposición excretas.....	XXX
Tabla 9.	Niveles de variable Techo.....	XXXI
Tabla 10.	Extracto información antropométrica	54
Tabla 11.	Extracto percentiles talla para la edad en niñas OMS	55
Tabla 12.	Extracto percentiles talla para la edad en niños OMS	56
Tabla 13.	Extracto percentiles peso para la talla en niños OMS	57
Tabla 14.	Extracto percentiles peso para la talla en niñas OMS	58
Tabla 15.	Clasificación de Waterlow para los casos estudiados	59
Tabla 16.	Desviaciones estándar de talla para la edad en niños según OMS.....	60
Tabla 17.	Desviaciones estándar de talla para la edad en niñas según OMS.....	60
Tabla 18.	Desviaciones estándar de peso para la talla en niños	61
Tabla 19.	Desviaciones estándar de peso para la talla en niñas	61
Tabla 20.	Clasificación nutricional infantil según desviaciones estándar	62
Tabla 21.	Resumen de diagnóstico desviaciones estándar	62
Tabla 22.	Codificación de variables por niveles binomiales.....	63
Tabla 23.	Codificación dicotómica de resultados de diagnóstico Waterlow	65
Tabla 24.	Codificación dicotómica desviaciones estándar (ZTE y ZPT)	65
Tabla 25.	Extracto codificación dicotómica para análisis de devianza.....	66
Tabla 26.	Análisis de devianza variable ZTE	67
Tabla 27.	Análisis de devianza variable ZPT	68
Tabla 28.	Análisis de devianza variable W	69
Tabla 29.	Resumen de variables significativas	71

Tabla 30.	Bloque 0: modelo ZTE	72
Tabla 31.	Variables en la ecuación modelo ZTE (bloque inicial)	72
Tabla 32.	Variables que no están en la ecuación modelo ZTE.....	73
Tabla 33.	Pruebas omnibus sobre los coeficientes del modelo ZTE	73
Tabla 34.	Prueba de Hosmer y Lemeshow del modelo ZTE	73
Tabla 35.	Tabla de contingencias para la prueba de Hosmer y Lemeshow	74
Tabla 36.	Tabla de clasificación modelo ZTE	74
Tabla 37.	Variables en la ecuación modelo ZTE	75
Tabla 38.	Bloque 0: modelo ZPT	76
Tabla 39.	Variables en la ecuación modelo ZPT (bloque inicial)	76
Tabla 40.	Variables que no están en la ecuación modelo ZPT.....	76
Tabla 41.	Pruebas ómnibus sobre los coeficientes del modelo ZPT	77
Tabla 42.	Prueba de Hosmer y Lemeshow del modelo ZPT	77
Tabla 43.	Tabla de contingencias para la prueba de Hosmer y Lemeshow	77
Tabla 44.	Tabla de clasificación modelo ZPT	78
Tabla 45.	Variables en la ecuación modelo ZPT	78
Tabla 46.	Bloque 0: modelo W	79
Tabla 47.	Variables en la ecuación modelo W (bloque inicial).....	79
Tabla 48.	Variables que no están en la ecuación modelo W	80
Tabla 49.	Pruebas omnibus sobre los coeficientes del modelo W.....	80
Tabla 50.	Prueba de Hosmer y Lemeshow del modelo W.....	80
Tabla 51.	Tabla de contingencias para la prueba de Hosmer y Lemeshow	81
Tabla 52.	Tabla de clasificación modelo W	81
Tabla 53.	Variables en la ecuación modelo W.....	82
Tabla 54.	Factores y niveles con efecto positivo variable ZTE	83
Tabla 55.	Factores y niveles con efecto positivo variable ZPT	84

Tabla 56.	Factores y niveles con efecto positivo variable W	85
Tabla 57.	Análisis de criticidad relativa de variables independientes	89
Tabla 58.	Análisis de devianza modelo ZTE.....	90
Tabla 59.	Análisis de devianza específica modelo ZPT.....	91
Tabla 60.	Análisis de devianza específica modelo W	91
Tabla 61.	Verosimilitud según prueba de chi cuadrado modelo ZTE.....	93
Tabla 62.	Verosimilitud según prueba de chi cuadrado modelo ZPT.....	93
Tabla 63.	Verosimilitud según prueba de chi cuadrado modelo W	94
Tabla 64.	Análisis del estadístico de Wald	95
Tabla 65.	Validación de modelos de regresión según AIC	96
Tabla 66.	Validación de modelos prueba de Hosmer Lemeshow	96
Tabla 67.	Cambio en probabilidades con niveles de efecto positivo.....	98

LISTA DE SÍMBOLOS

Símbolo	Significado
e^{β_i}	Antilogaritmo de coeficientes beta
R^2	Coefficiente de determinación
$D(y, \mu)$	Devianza
X^2	Estadístico de chi cuadrado para bondad de ajuste
Z	Estadístico de distribución normal
d	Estadístico de Durbin Watson
T	Estadístico T-Student
Wald	Estadístico de Wald para la bondad de ajuste
IMC	Índice de masa corporal
$\text{logit}(p_i)$	Logaritmo de probabilidades binomiales de respuesta
$\ln(\hat{\mu}; y)$	Logaritmo de verosimilitud del modelo de interés
$\ln(\hat{\mu}; \mu)$	Logaritmo de verosimilitud del modelo saturado

K	Número de parámetros estimado para un modelo de regresión múltiple
$P(ZTE)$	Probabilidad de que el individuo esté en condición de desnutrición según modelo de talla para la edad
$P(ZPT)$	Probabilidad de que el individuo esté en condición de desnutrición según modelo de peso para la talla
$P(W)$	Probabilidad de que el individuo esté en condición de desnutrición según modelo Waterlow
PY_i	Probabilidad de respuesta en modelos de regresión múltiple
$B_n X_n$	Término de ecuación de regresión múltiple conformado por el coeficiente y la variable independiente
\hat{u}^2	Valor de chi cuadrado para la prueba de Breusch-Pagan
L	Verosimilitud del modelo

GLOSARIO

AIC	Criterio de información de Akaike.
Antropometría	Es la ciencia que consiste en poder identificar a una persona con la ayuda de medidas y signos particulares individuales.
Chi cuadrado	Estadístico que mide la diferencia entre el modelo de regresión en su estado actual y el modelo cuando solo se incluyó una constante.
DA	Desnutrición aguda. Emaciación o deficiencia de peso para altura.
DCAG	Desnutrición crónica agudizada. Talla para la edad alterada y peso para la talla baja.
DCAR	Desnutrición crónica armonizada. Talla para la edad alterada y peso para la talla normal.
Desnutrición	Ingesta de calorías y nutrientes menor a la necesaria.
Devianza	Medida de bondad de ajuste que compara el modelo saturado con el modelo de interés.

Emaciación	Forma de malnutrición potencialmente mortal que provoca una delgadez y debilidad extremas en los niños, y aumenta sus posibilidades de morir o de sufrir deficiencias en su crecimiento, su desarrollo y su capacidad de aprendizaje.
Estadístico Wald	La prueba de Wald es un contraste de hipótesis donde se trata de ver la coherencia de afirmar un valor concreto de un parámetro de un modelo probabilístico una vez tenemos ya un modelo previamente seleccionado y ajustado.
IBM SPSS	Plataforma de software estadístico para extracción y procesamiento de datos.
Malnutrición	Carencia, exceso y desequilibrio de ingesta calórica y de nutrientes de una persona.
Odds Ratio	Relación de probabilidades.
OMS	Organización mundial de la salud.
OR	Odds Ratio.
Parsimonia	Principio que indica que la mejor explicación a un fenómeno es la más sencilla.

Regresión Logística	Análisis de regresión utilizado para predecir el resultado de una variable categórica en función de las variables dependientes o predictoras.
Retraso de crecimiento	Es un aumento de peso o estatura insuficiente o anormalmente lento en un niño menor de 5 años.
RL	Regresión Logística.
RSTUDIO	Entorno de desarrollo integrado para el lenguaje estadístico de programación R.
W	Clasificación de desnutrición según método de Waterlow.
ZPT	Clasificación de desnutrición según peso para la talla.
ZTE	Clasificación de desnutrición según talla para la edad.

RESUMEN

El propósito de la presente investigación fue desarrollar un modelo multivariado de regresión logística binomial que permitiera estudiar y explicar el comportamiento de las condiciones de desnutrición infantil en función de su relación a las variables físicas y energéticas de las viviendas en el área rural del suroccidente de Guatemala, para poder planificar estrategias de mejora con mayor eficiencia desde un enfoque centralizado en los factores que sean determinantes en esta problemática.

El objetivo general consistió en el desarrollo de un modelo estadístico multivariado utilizando la técnica de regresión logística binomial, para poder obtener probabilidades de presencia de la condición o evento de interés, en función de los niveles de los factores considerados.

El enfoque del estudio realizado es cuantitativo, ya que se desarrolló una herramienta que permitiera el cálculo de probabilidades de presencia de una condición en función de un conjunto de factores. El diseño de la investigación fue no experimental de tipo descriptiva, utilizando información retrospectiva. La metodología utilizada fue cuantitativa, calculando el tamaño de muestra con muestreo proporcional.

Se obtuvieron como resultados la magnitud de la relación logística entre las variables independientes y las condiciones de desnutrición según los métodos de diagnóstico utilizados. Asimismo, se desarrollaron los modelos de regresión logística con los que se identificaron los factores determinantes y en qué niveles

es más probable esperar condiciones de desnutrición, con sus probabilidades correspondientes.

Al analizar la información del presente estudio, se lograron desarrollar modelos de regresión logística para cada variable de respuesta, según los métodos de diagnóstico utilizados. Asimismo, se recomendó el uso de estos modelos para la planificación y ejecución de propuestas de mejora en el área rural del suroccidente de Guatemala, relacionadas a la problemática de desnutrición infantil.

PLANTEAMIENTO DEL PROBLEMA

- Contexto general

Durante varios años, el área rural de Guatemala ha sido un foco de casos de desnutrición infantil crónica entre toda la región Latinoamericana. De acuerdo con reportes de entidades nacionales e internacionales como UNICEF y SESAN, con un incremento del 16,4 % en casos de desnutrición aguda en menores de 5 años entre los años 2020 y 2021, y con estimaciones de 92 casos por cada 10 mil niños. Esta problemática se ha abordado con iniciativas desde diferentes perspectivas, que incluyen la reducción de pobreza y la mejora de obtención de alimentos, con poca mejoría en los resultados obtenidos, por lo que resulta necesario un fundamento estadístico para respaldar las mismas y direccionar de mejor forma los esfuerzos y recursos.

Desde el punto de vista estadístico, resulta necesario un modelo de regresión multivariado aplicado a las condiciones de desnutrición infantil, para mejorar las propuestas de desarrollo social en el suroccidente de Guatemala. Los programas que existen en la actualidad no consideran los principios de infraestructura de vivienda e higiene, definidos por la Organización Mundial de la Salud, para tener una vida saludable. Asimismo, se desconoce la medida en que estos principios se relacionan a las condiciones de desnutrición infantil a nivel nacional y específicamente en la región del suroccidente del país.

- Descripción del problema

No se ha desarrollado un modelo de regresión multivariado aplicado a las condiciones de desnutrición infantil, para mejorar las propuestas de desarrollo social en el suroccidente de Guatemala.

- Problemas específicos

- No se ha determinado cuál es la medida en la que se relacionan estadísticamente las condiciones de desnutrición infantil con las características físicas y energéticas de los hogares, para mejorar las propuestas de desarrollo social mediante el uso análisis de devianza en el suroccidente de Guatemala.
- No se ha explicado cuál es el comportamiento de las condiciones de desnutrición infantil, en relación con la información correspondiente a los hogares, para mejorar las propuestas de desarrollo social por medio de un modelo de regresión logística multivariada, en el suroccidente de Guatemala.
- No se ha han determinado cuáles son los factores físicos correspondientes a los hogares que son significativos en las condiciones de desnutrición infantil, y en qué niveles estos factores provocan condiciones de desnutrición, para mejorar las propuestas de desarrollo social mediante un modelo de regresión logística multivariada en el suroccidente de Guatemala.

- Formulación del problema

Para resolver los problemas descritos, el presente estudio pretende responder a las siguientes preguntas de investigación:

- ¿Cómo se puede explicar el comportamiento de las condiciones de desnutrición infantil, y mejorar estadísticamente las propuestas de desarrollo social en el suroccidente de Guatemala?
- ¿Cuál es la medida en la que se relacionan estadísticamente las condiciones de desnutrición infantil con las características físicas y energéticas de los hogares, para mejorar las propuestas de desarrollo social mediante el uso de análisis de devianza en el suroccidente de Guatemala?
- ¿Cuál es el comportamiento de las condiciones de desnutrición infantil, en relación con la información correspondiente a los hogares, para mejorar las propuestas de desarrollo social por medio de un modelo de regresión logística binomial en el suroccidente de Guatemala?
- ¿Cuáles son los factores físicos correspondientes a los hogares que son significativos en las condiciones de desnutrición infantil, y en qué niveles estos factores provocan condiciones de desnutrición, para mejorar las propuestas de desarrollo social mediante un modelo de regresión logística multivariada y el análisis de prevalencia factorial en el suroccidente de Guatemala?

- Delimitación del problema

Para el estudio propuesto, se utilizaron datos de familias que residen en el área rural del suroccidente de Guatemala, recopilados desde el año 2017 hasta el año 2021. La información pertenece a una organización de implementación de tecnologías agrícolas modernas en estas localidades, que, como parte de sus actividades normales, incluyen la realización de encuestas para recopilar datos económicos, estructurales y nutricionales.

La región en estudio comprendió un total 27 municipios del área rural del suroccidente de Guatemala, pertenecientes a los departamentos de Huehuetenango, Quiché y San Marcos, los cuales en su mayoría se dedican a actividades agrícolas.

OBJETIVOS

General

Desarrollar un modelo de regresión multivariado aplicado a las condiciones de desnutrición infantil, para mejorar las propuestas de desarrollo social en el suroccidente de Guatemala.

Específicos

1. Determinar la medida en la que se relacionan estadísticamente las condiciones de desnutrición infantil con las características físicas y energéticas de los hogares, para mejorar las propuestas de desarrollo social mediante el uso de análisis de devianza en el suroccidente de Guatemala.
2. Explicar el comportamiento de las condiciones de desnutrición infantil, en relación con la información correspondiente a los hogares, para mejorar las propuestas de desarrollo social por medio de un modelo de regresión logística multivariada en el suroccidente de Guatemala.
3. Determinar cuáles son los factores físicos correspondientes a los hogares que son significativos en las condiciones de desnutrición infantil, y en qué niveles estos factores provocan condiciones de desnutrición, para mejorar las propuestas de desarrollo social mediante un modelo de regresión logística multivariada y el análisis de *odds* ratios en el suroccidente de Guatemala.

RESUMEN DEL MARCO METODOLÓGICO

A continuación, se exponen los criterios adoptados para el proceso investigativo. Se describen las características principales del estudio, el planteamiento del problema, las variables de interés y las fases que se implementaron para el logro de los objetivos propuestos.

- Características del estudio

El enfoque de la investigación propuesta es cuantitativo, debido a que se busca una herramienta matemática que permita explicar el comportamiento de variables antropométricas numéricas relacionadas a condiciones nutricionales infantiles.

El diseño adoptado fue no experimental, puesto que las variables estudiadas no fueron manipuladas ni modificadas por condiciones externas. El diseño adoptado también se clasifica como retrospectivo, debido a que la toma de datos se realizó mientras estaba aconteciendo el fenómeno a estudiar (condiciones nutricionales). Asimismo, se clasifica como transversal, debido a que la toma de datos se realizó una vez para cada elemento de la muestra.

El tipo del estudio realizado es descriptivo, dado que se caracterizó y explicó estadísticamente el comportamiento de un fenómeno, de manera que se generara conocimiento útil para el desarrollo de propuestas de mejora a futuro.

- Variables del estudio

Las variables estudiadas durante el proceso de esta investigación fueron: peso, talla, materiales del techo, materiales del piso, el tipo de casa, fuente y calidad de agua, combustible de la cocina, disposición de basura y disposición de excretas, con el fin de obtener resultados para el análisis de la investigación. Las definiciones pueden observarse en la tabla 1.

Tabla 1.

Definición de variables

Variable	Definición Teórica	Definición Operativa
Presencia de desnutrición	Indica la presencia o ausencia de cualquiera de los tipos de desnutrición en un individuo, según la OMS.	Variable binaria en escala dicotómica que indica si el individuo analizado presenta una condición de desnutrición.
Material piso	Materia prima o producto empleado en la construcción del piso de una vivienda.	Variable categórica cualitativa nominal, que indica el material del que está elaborado el piso de la vivienda
Tipo casa	Clasificación estructural de las viviendas, según dimensión y características de construcción.	Variable categórica cualitativa nominal, que indica la clasificación de la casa según su estructura.
Fuente agua	Proveniente del agua utilizada en los hogares.	Variable categórica cualitativa nominal, que indica la clasificación de las fuentes de agua en las viviendas.
Calidad agua	Calificación del estado en que se utiliza el agua para consumo humano.	Variable categórica cualitativa nominal, que indica el estado del agua potable en las viviendas.

Continuación de la tabla 1.

Combustible cocina	Clasificación del tipo de combustible utilizado para elaborar alimentos.	Variable categórica cualitativa nominal, que indica el tipo de combustible usado en la cocina.
Disposición basura	Clasificación del método usado para extraer desechos domésticos.	Variable categórica cualitativa nominal, que indica la forma en que se retiran los desechos de la vivienda.
Disposiciones excretas	Clasificación del método usado para extraer desechos sanitarios.	Variable categórica cualitativa nominal, que indica la forma en que se retiran los desechos sanitarios.
Material de techo	Clasificación del material usado en la estructura del techo de la vivienda	Variable categórica cualitativa nominal, que indica la forma en que se retiran los desechos sanitarios.

Nota. Elaboración propia, realizado con Microsoft Word.

- Niveles de las variables del estudio

Al considerar que todas las variables independientes del estudio son categóricas, se presentan los niveles de clasificación designados para cada una en las siguientes tablas.

Tabla 2.

Niveles de variable Material piso

Variable	Nivel
Material piso	Tierra, madera y otro
Material piso	Ladrillo
Material piso	Torta de cemento
Material piso	Granito o ladrillo cerámico

Nota. Elaboración propia, realizado con Microsoft Word.

Tabla 3.

Niveles de variable Tipo casa

Variable	Nivel
Tipo casa	Palomar
Tipo casa	Rancho
Tipo casa	Vivienda improvisada
Tipo casa	Casa formal
Tipo casa	Apartamento
Tipo casa	Otro tipo de vivienda

Nota. Elaboración propia, realizado con Microsoft Word.

Tabla 4.

Niveles de variable Fuente agua

Variable	Nivel
Fuente agua	Chorro público
Fuente agua	Lluvia
Fuente agua	Pozo
Fuente agua	Río
Fuente agua	Camión o pipa
Fuente agua	De otra vivienda
Fuente agua	Tubería dentro de vivienda o terreno

Nota. Elaboración propia, realizado con Microsoft Word.

Tabla 5.

Niveles de variable Calidad agua

Variable	Nivel
Calidad agua	Filtrada
Calidad agua	Clorada
Calidad agua	Purificada
Calidad agua	Hervida
Calidad agua	Ninguna

Nota. Elaboración propia, realizado con Microsoft Word.

Tabla 6.

Niveles de variable Combustible cocina

Variable	Nivel
Combustible cocina	Leña
Combustible cocina	Gas
Combustible cocina	Estufa eléctrica

Nota. Elaboración propia, realizado con Microsoft Word.

Tabla 7.

Niveles de variable Disposición basura

Variable	Nivel
Disposición basura	Servicio formal
Disposición basura	Servicio informal

Nota. Elaboración propia, realizado con Microsoft Word.

Tabla 8.

Niveles de variable Disposición excretas

Variable	Nivel
Disposiciones excretas	Letrina
Disposiciones excretas	Inodoro
Disposiciones excretas	Intemperie

Nota. Elaboración propia, realizado con Microsoft Word.

Tabla 9.

Niveles de variable Techo

Variabl	Nivel
Techo	Lámina
Techo	Asbesto
Techo	Teja de barro
Techo	Concreto
Techo	Otro

Nota. Elaboración propia, realizado con Microsoft Word.

- **Muestreo**

Al considerar la población infantil para los departamentos de: Totonicapán, Quetzaltenango, San Marcos, Huehuetenango, Sololá y Quiché que conforman el altiplano suroccidental de Guatemala, se puede obtener una población de niños menores de 2 años equivalente a 591168, de acuerdo con los datos del censo poblacional realizado en Guatemala en el año 2018.

Para el plan de muestreo probabilístico con población finita, se considerará un error de 5 % y una confianza de 95 % y un porcentaje de familias que accederán a colaborar con el estudio en un 80 %, lo que indica que un 20 % no querrá colaborar o no estará presente.

Fórmula de muestreo para el cálculo de tamaño mínimo de muestra:

$$n = \frac{k^2 pqN}{e^2(N-1) + k^2 pq} \text{ [Ec. 1]}$$

En donde:

n: tamaño muestra a determinar

p: variabilidad positiva = (0.80)

q: variabilidad negativa = (0.20)

N: tamaño de la población = (591168)

e: es el error muestral que se considera = (0.05)

k: constante de nivel de confianza que indica la probabilidad de que los resultados del estudio sean ciertos o no. Sustituyendo los datos se obtiene con K = 95%:

$$n = \frac{(1.96^2)(0.5)(0.5)(591168)}{(0.05^2)(591168 - 1) + \{(1.96^2)(0.50)(0.5)\}} = 245.76 \approx 246$$

La muestra fue la siguiente:

- Población aproximada de niños menores a 2 años en el suroccidente de Guatemala: 591168.

- Muestra de niños para toma de datos: el mayor número posible que sobrepase el tamaño mínimo de muestra.
- Mínimo de niños menores a 2 años incluidos en las encuestas: 246.

Para los objetivos del estudio realizado también se considera una población infinita, como en muchos estudios de regresión logística. Por esta razón, se puede calcular un tamaño de muestra en función de la cantidad de variables predictoras utilizadas.

Para el cálculo de tamaño de muestra se utiliza la fórmula de Freeman de población infinita con un promedio de eventos por variables igual a 50:

$$n = 50 * (k + 1) \text{ [Ec. 2]}$$

En donde:

n: tamaño muestra a determinar

k: cantidad de predictores incluidos = 9

$$n = 50 * (9 + 1) = 500$$

Con los resultados de los 2 métodos de muestreo utilizados se considera el de tamaño de muestra de mayor cantidad para incrementar la exactitud estadística.

Al considerar que, al momento de diseño del presente estudio, se contaba con una cantidad de 1930 individuos con registros antropométricos digitales viables para el análisis de información, se consideró en su totalidad esta muestra para el análisis y mejorar la potencia estadística.

- Fases del estudio

Fase 1: revisión de literatura

Esta fase comprendió actividades de investigación y revisión de fuentes bibliográficas para los fundamentos teóricos y antecedentes relacionados a la problemática del estudio.

Fase 2: gestión o recolección de la información

Esta fase comprendió las actividades de acceso a la base de datos administrada por la organización facilitadora, la clasificación de niveles y factores, el cálculo de diagnóstico nutricional y transformación de variable respuesta.

Fase 3: análisis de información

En esta fase se realizaron las actividades de correlación múltiple, diseño de modelo de regresión lineal, validación de supuestos y evaluación de bondad de ajuste.

Fase 4: interpretación de información

En esta fase se estudiaron detenidamente los resultados obtenidos del modelo de regresión logística, se identificaron los factores significativos y se evaluó la exactitud del modelo.

INTRODUCCIÓN

El estudio realizado surge como una necesidad de innovación de información estadística descriptiva y correlativa por parte de entidades gubernamentales y no gubernamentales, que deseen implementar iniciativas para mejorar la problemática de desnutrición infantil en el área rural del suroccidente de Guatemala. El estudio fue realizado dentro de la línea de investigación de análisis multivariado y regresión.

El problema estudiado es la alta prevalencia de desnutrición infantil en Guatemala, según datos recientes reportados por organizaciones nacionales e internacionales (alrededor del 50% de los niños entre 0 y 5 años se encuentran en algún estado de desnutrición), y las estrategias de mejora que no han sido certeras ni eficientes durante los últimos años.

La realización del presente estudio resulta de importancia debido a la necesidad existente de generación de conocimiento que permita direccionar iniciativas de mejora, al saber el efecto que tienen factores relacionados a la infraestructura de viviendas y saneamiento en estas áreas.

Para obtener el conocimiento utilizado, se aplicaron metodologías estadísticas como el análisis de varianza multifactorial y la regresión logística, así como sus supuestos y técnicas evaluadoras de ajuste. El estudio realizado es de enfoque cualitativo con datos retrospectivos, de carácter no experimental y tipo descriptivo.

Con los resultados obtenidos durante el desarrollo de la presente investigación, se lograron desarrollar 3 modelos estadísticos de predicción de condiciones de desnutrición, en función de variables predictoras correspondientes a las condiciones de infraestructura de vivienda e higiene. Asimismo, se identificaron los valores de relación estadística entre estas variables y las condiciones de desnutrición, utilizando los modelos desarrollados para determinar los niveles de los factores con efectos positivos en las condiciones de desnutrición estudiadas. Con el conocimiento obtenido durante la realización del estudio, es posible mejorar el respaldo de las iniciativas de mejora que se planteen en el futuro por diversas organizaciones, así como ampliar y profundizar estadísticamente el conocimiento que se tiene relacionado a la problemática de desnutrición en áreas rurales de Guatemala y países con condiciones económicas y sociales similares.

El desarrollo del estudio se dividió en varias fases. La primera fase consistió en la revisión de literatura para el conocimiento referencial y conceptual. La segunda consiste en la gestión y recolección de la información mediante la consulta de la base de datos de las personas visitadas, la organización de datos y el cálculo de diagnóstico nutricional. La tercera fase en el análisis estadístico de la información recopilada, con las técnicas multivariadas utilizadas. Por último, en la quinta fase se discuten los resultados obtenidos, describiendo los hallazgos realizados, y haciendo la comparación con estudios previos revisados.

Para la realización del presente estudio fue necesaria una inversión aproximada de 40 mil quetzales guatemaltecos, utilizados para cubrir gastos de recursos humanos, recursos materiales, gastos de planificación y gastos de ejecución.

El informe se integra por cuatro capítulos. El primero presenta el marco referencial, donde se describen los conocimientos de estudios internacionales y nacionales sobre desnutrición infantil y regresión logística, que se encuentra presente en la mayoría de los estudios revisados.

El segundo capítulo describe el marco teórico donde se abordan los temas principales concernientes a los estudios multivariados en general y a la regresión logística binomial. Asimismo, se presenta la bibliografía correspondiente a los métodos de diagnóstico de desnutrición más utilizados y las variables a monitorear.

El tercer capítulo presenta los resultados obtenidos del análisis de devianza realizado a nivel de factores, y los resultados de los modelos de regresión desarrollados, incluyendo los Odds Ratios usados para calcular las probabilidades de presencia de desnutrición entre los niveles de las variables estudiadas.

El cuarto capítulo presenta la discusión de los resultados obtenidos, realizando la validación correspondiente de los modelos según las pruebas estadísticas de bondad de ajuste y complejidad aplicadas. Asimismo, se muestran las probabilidades relativas de las variables significativas relacionadas a la presencia de desnutrición calculadas con los odds ratios obtenidos y los contrastes con estudios anteriores.

Al finalizar el desarrollo del presente estudio se concluyó que las condiciones de desnutrición presentan relación estadística significativa con varias variables estudiadas, correspondientes a la infraestructura de vivienda e higiene, entre las que destacan la fuente de agua, el piso de la casa y la disposición de excretas. Con estas variables y la relación identificada, se lograron desarrollar modelos de regresión logística binomial, que pueden ser utilizados para probabilidades de presencia de desnutrición infantil. Asimismo, con los resultados de los modelos desarrollados se calcularon los odds ratios o razones de probabilidad, para determinar los niveles de variables significativas que ejercen un efecto positivamente peligroso en la presencia de desnutrición infantil. Se sugiere utilizar los hallazgos y modelos realizados para mejorar las propuestas de desarrollo en el marco de la desnutrición infantil en áreas rurales de Guatemala y países similares.

1. MARCO REFERENCIAL

El panorama nutricional infantil en Guatemala ha presentado problemas evidentes tanto de desnutrición crónica como desnutrición crónica severa, de acuerdo con diagnósticos sistemáticos que han sido realizados por varias instituciones desde el año 1995.

Guatemala ha mostrado históricamente niveles alarmantes de desnutrición crónica en niños menores de 5 años, lo que constituye una violación sistemática de sus derechos. Es una causa común de la baja talla, de daños estructurales en el cerebro, que produce retraso en el desarrollo de las funciones cognitivas de manera permanente y en el crecimiento, el cual a los 2 años se vincula con una reducción en la escolaridad de 0.9 años y un incremento del 16 % en el riesgo de reprobación al menos un grado en la escuela.

Si bien en los últimos años los datos muestran una reducción de la desnutrición crónica en menores de 5 años, esta variación no ha sido del todo alentadora, pues en un período de 20 años (1995-2015) disminuyó solamente 8.7 puntos porcentuales. Si se comparan los resultados de la VI Encuesta Nacional de Salud Materno Infantil (Ensmi) 2014/2015 elaborada por el Ministerio de Salud Pública y Asistencia Social (MSPAS),

el Instituto Nacional de Estadística (INE) y (Segeplan) (2017), con la V Ensmi 2008/2009, la desnutrición crónica en menores de 5 años registró una reducción de apenas 3 puntos porcentuales en un período de 6 años al pasar de 49.8 % a 46.5 %, con promedios anuales de reducción de alrededor de ½ punto porcentual. En cuanto a la desnutrición crónica severa, de 2009 a 2015 registró una reducción levemente más acentuada que en el caso de la desnutrición crónica total. (ICEFI, 2019, p. 7)

En años recientes, se han elaborado diversos estudios que implementan los métodos estadísticos de análisis multivariado, en los que se han incorporado variables relacionadas a la nutrición para explicar el comportamiento de los índices antropométricos, en varios conjuntos poblacionales. Entre los grupos de características y factores que se adoptan como variables independientes, se pueden mencionar las condiciones socioeconómicas, el conocimiento de la dieta nutritiva para los menores de 5 años, las características nutricionales del calostro y del lactante, las condiciones de higiene y seguridad de la vivienda, entre otras.

Los resultados de las investigaciones y estudios mencionados son variados, y le atribuyen mayor o menor efecto sobre la variable respuesta (indicadores antropométricos y otros indicadores nutricionales), según la población de estudio, la naturaleza de la investigación, y los métodos implementados.

El estudio de indicadores nutricionales y la aplicación de técnicas multivariadas para la explicación de su comportamiento, ha sido realizado en varios países del mundo y con mayor frecuencia en Latinoamérica, con diversos

objetivos. Sin embargo, la mayoría de ellos se relacionan con mejorar la situación nutricional infantil debido a que esta región presenta una problemática significativa en este aspecto.

López (2015) realizó un estudio para explicar los patrones de crecimiento craneofacial en niños de cero a cuatro años residentes en Bogotá, Colombia. Las variables dependientes consideradas fueron el peso, la talla y dimensiones craneofaciales.

Los métodos de análisis de datos fueron el análisis de varianza, análisis de componentes principales, cadenas de Márkov y el modelo multivariado de estimación bayesiana, con el que se construyeron curvas de crecimiento no lineales. Las variables explicativas de mayor asociación fueron: género, ciudad de nacimiento de los padres y abuelos, lactancia materna, estrato socioeconómico, nivel de ingresos, nivel de escolaridad de los padres, tiempo de gestación, condiciones de peso y talla al nacer, servicios públicos, tipo de vivienda, entre otros.

Valdivia et al. (2017), realizaron un estudio longitudinal cuyo objetivo era determinar la correlación entre las características proteicas del calostro y la leche materna con los indicadores antropométricos de la madre y del lactante, con el uso de información recopilada en un hospital de Huancayo, Perú. Los métodos estadísticos implementados fueron el análisis bivariado y la regresión logística multivariada. Los resultados obtenidos del análisis bivariado evidenciaron correlación significativa entre la concentración de proteínas del calostro y peso al cuarto mes y la concentración de proteínas de leche madura con respecto a la talla del lactante al cuarto mes, mientras que la regresión logística demostró relación significativa entre el peso del recién nacido y el índice de masa corporal pregestacional.

González y Vega (1994), realizaron un estudio para determinar la proporción de niños menores de un año que no alcanzaban el peso mínimo adecuado para su edad e identificar los factores que pudieran estar asociados con esa problemática, en áreas periféricas de la ciudad de Guadalajara, México. Mediante el método de regresión logística se lograron determinar los Odds Ratios, para la condición de que los niños del estudio que no alcanzarán el peso mínimo adecuado.

Los factores que se evidenciaron como significativos para los niños que no alcanzaban el peso mínimo fueron la edad de la madre, y la pertenencia materna a cierto grupo social. Los Odds Ratios mayores se presentaron cuando la edad de la madre era menor a 20 años, y el grupo social de pertenencia materna era obrero, según la clasificación designada para el estudio.

Por otra parte, Santiago et al. (2015), analizaron variables antropométricas en adultos jóvenes, dentro de escuelas públicas de Brasil, con exceso de peso para determinar si existe relación con características sociodemográficas. Se realizó un estudio de casos y controles, con el método de regresión logística múltiple. Los resultados evidenciaron, después de las etapas de ajuste de los modelos, que los factores significativos para la condición de obesidad eran el estado civil con compañero, aumento de peso en la adolescencia, presión arterial diastólica aumentada y circunferencia abdominal aumentada. Para validar el ajuste del modelo final según la regresión logística se utilizaron los coeficientes de determinación de Hosmer, Cox & Snell, y Nagelkerke.

En el estudio realizado por Álvarez (2019), se analizó la prevalencia de la desnutrición por antropometría, según lo establecido por la OMS, en 243 niños menores de 3 años en la provincia de Huánuco, Perú. Se calcularon estadísticos

descriptivos, medidas de asociación Chi cuadrado de Pearson, odds ratio y análisis multivariado (regresión logística binomial). Para la validación se utilizó el R cuadrado de Nagelkerke.

Los factores significativos según la regresión logística fueron el desconocimiento del número de raciones que debe consumir diariamente el niño a partir de los 6 meses, la inaccesibilidad económica para el abastecimiento de alimentos, el bajo apoyo social, la rotación de cultivos agrícolas tradicionales por aquellos comerciales ineficaces, el bajo peso al nacer, la baja escolaridad de la madre, la desnutrición materna durante el embarazo, el inadecuado saneamiento básico de la vivienda (desagüe, eliminación de residuos sólidos, agua, y energía eléctrica), la baja producción ante sequías, heladas, ranchas, entre otros.

Al continuar en el marco de las investigaciones para el estudio de las variables antropométricas, Lera et al. (2007), analizaron datos antropométricos longitudinales de una muestra de 57 mujeres adultas mayores de 60 años, durante un periodo de 30 meses en El Gran Santiago, Chile. El método estadístico fue el análisis de similaridad STATIS, que consiste en el desarrollo matricial de variables dependientes e independientes, de manera que sea posible generar una matriz de compromiso que resume las matrices longitudinales, y así obtener trayectorias en función del tiempo, que se utilizó como factor explicativo. Los resultados del análisis evidenciaron que las variables peso, circunferencia de brazo, pantorrilla, cintura y cadera, resultaron ser las más correlacionadas con el primer eje de compromiso. Las variables: talla y altura de rodilla, responden al tamaño del esqueleto. Este eje presenta gran estabilidad, debido a que la altura de la rodilla no cambia en el tiempo y en 30 meses la disminución de la talla es mínima.

Respecto a los estudios realizados en Guatemala correspondientes a condiciones nutricionales infantiles, existen varios de carácter diagnóstico y de monitoreo/evaluación, así como de clasificación de condiciones nutricionales, mayormente dentro del área rural, que describen en términos generales las condiciones socioeconómicas de la muestra incluida.

Una de las investigaciones a nivel nacional con mayor respaldo estadístico correlacional, es la realizada por Kragel et al. (2020), que consistía en el análisis de factores de riesgo relacionados al retraso del crecimiento en niños menores de 5 años, localizados en el altiplano rural de Guatemala. Los métodos de análisis fueron los estudios de prevalencia y ANOVA multifactorial.

En el estudio realizado por López Bonilla (2016), se analizó la prevalencia de desnutrición crónica, general y aguda, dentro de un grupo hospitalario de 322 niños mayores de un mes y menores de 5 años, en el departamento de pediatría del Hospital Nacional de Enfermedades ubicado en la ciudad de Guatemala. De acuerdo con la clasificación por percentiles establecida por la OMS para el análisis antropométrico, el 14.28 % de los niños del estudio se clasificaron con peso bajo y el 8.38 % con desnutrición aguda moderada. Como causas que producen estados de desnutrición en este medio, se atribuye el 90 % a la subalimentación, deficiencia de alimentos consumidos y falta de higiene en la vivienda.

Por otra parte, Castillo (2017), realizó un estudio descriptivo retrospectivo, para poder proponer mejoras presupuestales en el Centro de Recuperación Nutricional del área Ixil del Departamento de Quiché, Guatemala. Las principales causas de desnutrición mencionadas en el estudio incluyen factores socioambientales como el nivel de pobreza y pobreza extrema, la disponibilidad, acceso y consumo de alimentos, agua potable, sanidad, educación, y

saneamiento ambiental. Se menciona que en el área Ixil existen problemas de contaminación de agua y sanidad de viviendas. Como criterios de inclusión se seleccionaron niños entre 6 meses y 5 años con diagnóstico de desnutrición aguda moderada y severa, marasmo, o diagnóstico de Kwashiorkor.

En el estudio transversal realizado por Alvarado et al. (2013), se describieron las condiciones sociales y familiares en municipios del departamento de Chimaltenango, Guatemala, para explicar las diferencias en prevalencia de desnutrición crónica en menores de 5 años. Entre los factores analizados se incluyó la cantidad de hijos, la presencia de ambos padres, la actividad económica y escolaridad de los padres, y las condiciones higiénicas y de infraestructura de vivienda. Respecto a la vivienda, el 69.4 % contaban con agua intradomiciliaria, mientras que el 42.1 % tenían letrina, 33.3 % sanitario, y 52 % piso de tierra.

En el estudio publicado por la ASB (2016), cuyo objetivo era conocer el estado nutricional de niñas y niños menores de 5 años y el grado de (in)seguridad alimentaria de sus hogares en el área rural en municipios guatemaltecos de Zacapa y Chiquimula, se analizaron datos cuantitativos y transversales de 465 familias. Entre los resultados obtenidos, se puede destacar la prevalencia de desnutrición aguda global, equivalente a un 3.8 % y una prevalencia de desnutrición crónica del 72.1 %, así como un 33.3 % de prevalencia en desnutrición global (bajo peso). Entre la información descriptiva relacionada a la infraestructura de vivienda destaca un 93.3 % de viviendas construidas con adobe o caña/barro y un 84.7 % de viviendas con piso de tierra, y un poco más de la mitad de las familias estudiadas tiene disposición de excretas a la intemperie.

Con la información obtenida al investigar y revisar los estudios correlacionales de condiciones nutricionales descritos previamente, tanto en ámbitos nacionales como internacionales, se establecieron directrices, estrategias y métodos a aplicar en el presente estudio. Al considerar que la mayoría de los estudios utilizaron la regresión logística para identificar la magnitud de las variables explicativas con las condiciones de desnutrición, se optó por usar esta técnica en el presente estudio.

Gracias a la información revisada en los estudios previos, se consideraron las variables de infraestructura de vivienda, higiene y sanitización como explicativas para el presente estudio. Asimismo, al considerar que gran parte de los estudios consultados utilizaron la medición de variables antropométricas para la identificación y clasificación de casos de desnutrición, se optó por aplicar los mismos procedimientos de diagnóstico en el presente estudio, cuya base son las tablas generalizadas de crecimiento publicadas por la OMS.

2. MARCO TEÓRICO

Esta sección está dirigida a presentar la base teórica para dar a conocer el fundamento del conocimiento en materia de temas estadísticos referentes a los métodos de análisis multivariado aplicado al análisis de datos. Asimismo, se describen los fundamentos teóricos relacionados a los índices nutricionales infantiles.

2.1. Fundamentos estadísticos

Entre los fundamentos estadísticos a utilizar en el estudio, se considera la regresión múltiple con sus supuestos, que forma parte de la estadística multivariada, y específicamente la regresión logística multivariada.

2.1.1. Regresión múltiple

Las variables dependientes en ocasiones se pueden explicar de mejor forma al utilizar un modelo multivariado, cuando existe relación con muchas variables predictoras.

En muchas ocasiones una sola variable independiente es insuficiente para predecir el comportamiento de una variable dependiente, por lo que necesita del aporte de otras variables independientes y cuando se crea un modelo, a partir de un conjunto de variables para poder estimar el valor de una variable respuesta es cuando se aplica la regresión múltiple, que

es el resultado de diversas variables y su efecto en una sola variable dependiente. (Levine & Rubin, 2004, p. 582)

Todos los modelos de regresión múltiple creados a partir de una base de datos, pueden representarse por medio de la siguiente ecuación general:

$$Y_i = B_0 + B_1X_1 + B_2X_2 + \dots + B_nX_n + \epsilon \text{ [Ec. 3]}$$

En dicha ecuación, viene dado por una cantidad de variables independientes que a su vez están acompañadas de un coeficiente, el cual indica el peso o el grado de influencia que cada una de las variables independientes tiene sobre el valor Y, además de contar con un valor ϵ y X_n . En el cual el primero representa una variable aleatoria y el segundo representa los errores, ambas variables son utilizadas para representar lo que los coeficientes de las variables independientes no son capaces de explicar dentro del modelo. (Levine & Rubin, 2004, p. 582)

2.1.2. Supuestos estadísticos de la regresión múltiple

Para realizar la regresión múltiple es necesario validar algunos supuestos estadísticos antes y después del proceso de análisis de datos.

Los supuestos de la regresión lineal múltiple son aquellas pruebas que permiten determinar la validez del modelo creado, al analizar desde sus coeficientes hasta sus residuos. Con el cumplimiento de estos supuestos es posible determinar si el modelo es predictor de la variable dependiente o no. (Valle, 2018, p. 14)

2.1.2.1. Independencia

La independencia es un supuesto aplicado a los residuos resultantes de un modelo de regresión, que se basa en el principio de que los residuos deben presentar un comportamiento o patrón evidente; esto indicaría que el modelo estudiado no se desempeña igual para todos los rangos de valores porque los residuos dependen unos de otros.

Este supuesto se refiere al comportamiento que deben presentar los residuos dentro de un modelo estadístico de regresión lineal múltiple, los cuales deben presentar total independencia entre ellos, es decir que, no sean producto de una tendencia dentro del modelo desarrollado. (Spiegel, 2009, p. 159).

La prueba de independencia sirve para establecer la existencia de alguna colinealidad entre las variables del modelo estudiado. La independencia hace referencia a que todas las variables que se pretende incluir en el modelo como una variable predictora, no dependan de ninguna otra

variable incluida en el modelo de regresión, porque de lo contrario, dicha variable no podrá ser utilizada. Cuando una variable independiente tiene alguna relación con cualquier otra del modelo o depende del comportamiento de otra, se dice que dichas variables presentan colinealidad y dicho fenómeno debe ser evitado en un modelo de regresión lineal múltiple. (Palomo, 2011, p. 46).

Para validar este supuesto se utiliza el estadístico de Durbin Watson, que se basa en obtener varias observaciones de errores, para hacer una sumatoria de cuadrados de la resta entre una observación y otra:

$$d = \frac{\overline{\sum_{t=2}^T (e_t - e_{t-1})^2}}{\sum_{t=1}^T e_t^2} \text{ [Ec. 4]}$$

T es el número de observaciones. Puesto que d es aproximadamente igual a $2(1 - r)$, donde r es el coeficiente de autocorrelación de primer orden de los residuos, d = 2 indica que no hay autocorrelación. El valor de d siempre está entre 0 y 4. Si el estadístico de Durbin-Watson es sustancialmente menor que 2, hay evidencia de correlación seria positiva.

2.1.2.2. Homocedasticidad

La homocedasticidad se refiere a que el comportamiento de residuos respecto a la varianza debe ser similar, para descartar la posibilidad de fallos dentro del modelo para ciertos rangos de datos. Es decir, que la varianza de los residuos debe ser constante.

La homocedasticidad del modelo multivariado plantea la hipótesis estadística que la variación de los residuos será independiente de las variables explicativas o independientes del modelo multivariado. Para la determinación de la homocedasticidad estadística de un modelo multivariado se graficaron los valores pronosticados del modelo contra el cuadrado de los residuos para buscar la independencia entre estas variables por medio de la gráfica. (Islas, 2010, p. 10)

Para comprobar el supuesto de homocedasticidad es común utilizar la prueba de Breusch-Pagan, en la que se analiza si la varianza estimada de los residuos de una regresión depende de los valores de las variables independientes, utilizando la distribución chi cuadrado:

$$\hat{u}^2 = \gamma_0 + \gamma_1 x + v \text{ [Ec. 5]}$$

Si el valor de chi cuadrado calculado en la prueba es menor al valor p definido, se rechaza la hipótesis nula de homocedasticidad, y se acepta la hipótesis de heterocedasticidad.

2.1.2.3. Multicolinealidad

Como un supuesto enfocado en la reducción de la complejidad de los modelos, la multicolinealidad mide si las variables independientes presentan correlación. En caso se identifique correlación entre ellas, se deben eliminar del modelo para reducir el ruido y la varianza de los coeficientes estimadores que pertenecen a la ecuación de regresión resultante del modelo.

Este supuesto también es conocido como no-colinealidad y se refiere a que dentro de las variables independientes del modelo de regresión lineal no debe existir ningún tipo de correlación entre ellas, es decir, una variable independiente no dependa de otra u otras variables independientes. (Levine & Rubin, 2004, p. 588)

La multicolinealidad puede ser de dos tipos:

- Multicolinealidad perfecta: una de las variables independientes tiene una regresión lineal con una o más variables independientes.
- Multicolinealidad parcial: este tipo de colinealidad resulta cuando existe una alta correlación entre las variables independientes del modelo de regresión.

La multicolinealidad se aprecia cuando las variables independientes del modelo desarrollado se expresan como una combinación lineal:

$$0 = B_0 + B_1X_1 + B_2X_2 + \dots + B_nX_n \text{ [Ec. 6]}$$

Una de las formas más sencillas para determinar la colinealidad entre variables independientes es con la determinación del coeficiente de determinación R^2 para cada variable comparada con las demás. Al ser detectada cualquier dependencia entre las variables independientes, éstas deben ser eliminadas del modelo creado, porque de no ser así esta

colinealidad afectará en incrementar la varianza de los coeficientes estimadores de la ecuación. (Kutner, Nachtsheim, Neter, & Li, 2005, p. 407).

Para calcular el coeficiente de determinación se utiliza la covarianza de las 2 variables que se desean comparar dividida entre la varianza de cada una:

$$R^2 = \frac{\sigma^2_{XY}}{\sigma^2_X \sigma^2_Y} \text{ [Ec. 7]}$$

2.1.2.4. Prueba de hipótesis

Con el propósito de validar si los parámetros estadísticos cumplen o no con alguno de los supuestos planteados para la técnica estadística a utilizar, como la regresión multivariada, se utilizan pruebas de hipótesis. Las pruebas de hipótesis comparan valores teóricos seleccionados de los datos de acuerdo con alguna distribución estadística con valores observados o estadísticos de prueba, para aceptar o rechazar hipótesis o proposiciones que se hayan definido para un conjunto de datos.

En el amplio campo de la investigación, en muchas ocasiones se vuelve vital contar con un criterio para aceptar o rechazar algún supuesto planteado sobre alguno de los parámetros estadísticos, este supuesto recibe el nombre de “hipótesis estadística” que se ha convertido en una de las herramientas de mayor peso en la inferencia estadística, porque permite tomar una posición fundamentada al momento de concluir sobre

las características de un modelo. A través de las llamadas “Pruebas de hipótesis”. (Levine & Rubin, 2004, p. 320)

En estadística se aplican diferentes tipos de pruebas de hipótesis para estimar las características de una distribución, hipótesis nula y la hipótesis alternativa. La primera es la que regularmente se plantea para alguno de los parámetros que se desee evaluar, mientras que la hipótesis alternativa es aquella que contradice lo que la hipótesis nula plantea, es decir, que la hipótesis alternativa pasará a ser aceptada al momento de rechazar la hipótesis nula que se planteó al inicio de la prueba.

Para la determinación de las conclusiones de cada una de las hipótesis propuestas en los supuestos, las pruebas de hipótesis son fundamentadas en distribuciones teóricas en las cuales se comparan dos valores, uno calculado con los valores reales de la distribución creada y otro valor teórico buscado en las diferentes tablas, según sea la prueba que se aplique. (Valle, 2018, p. 18)

Las distribuciones más utilizadas para realizar pruebas de hipótesis son las siguientes: distribución T-Student y distribución normal (Z).

- T-Student

La prueba T-Student se utiliza para contrastar hipótesis respectivas a medias cuando los datos no necesariamente siguen la distribución normal, de acuerdo con la cantidad de grados de libertad disponibles, los cuales dependen de la cantidad de datos de la muestra o población.

La prueba T-Student fue creada en 1899, por William Sealey Gosset, químico inglés que se dedicaba al control de calidad, quien publicó sus hallazgos en el tema de estadística anónimamente con el nombre de "Student". Esta prueba es utilizada para distribuciones que atienden a la normalidad y en la cual se pueden contrastar hipótesis sobre medias, también puede ser utilizada para distribuciones que no son precisamente normales cuando se trabaja con grandes poblaciones, sin embargo, no es recomendable usar esta prueba para esos casos. Esta distribución está determinada a ser usada para muestras pequeñas ($n < 30$).

Para la distribución de T-Student es importante establecer que la curva en la cual se distribuyen los datos depende de un factor importante, los grados de libertad, los cuales corresponden a un número entero positivo que representa la cantidad de variables que están libremente determinadas bajo el hecho que la suma de las diferencias entre las observaciones y el valor medio debe dar como resultado siempre cero.

La curva de la distribución T-Student varía, según la cantidad de grados de libertad que presente la serie de datos, pero ésta no se aleja de la similitud con una campana Gaussiana, también cuenta con el centro o valor medio cero y una dispersión simétrica hacia los lados; sin embargo, esta dispersión de los datos es mayor con una cantidad de grados de libertad bajo. A medida que los grados de libertad de la distribución aumenten, la campana de distribución T-Student se asemejará más a una distribución normal estándar.

Para la evaluación de pruebas de hipótesis con la distribución T-Student se hace una comparación entre dos valores de T dados por las especificaciones de un valor alfa (α) y los grados de libertad calculados para la serie de datos, valor que es determinado por tablas ya creadas para diferentes niveles de confianza y grados de libertad que trabaje la distribución analizada, con un valor de T calculado a partir de la siguiente ecuación:

$$T = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \text{ [Ec. 8]}$$

En dependencia de los criterios que la prueba de T-Student utilizada plantee, la hipótesis nula propuesta puede ser aceptada o rechazada,

según sean los valores para la T teórica y el valor de la T calculada.

(Noste, 2013, p. 102)

- Distribución normal

La distribución normal es ampliamente utilizada para contrastar hipótesis realizadas a medias de conjuntos de datos cuyo tamaño de muestra es grande. Posee la característica de convertir cualquier valor real en valores Z que se pueden comparar con las tablas de la campana de Gauss para aceptar o rechazar hipótesis.

La distribución fue estudiada por el matemático Gauss, la cual se trata de una variable aleatoria continua que puede tomar cualquier valor, siempre y cuando éste se encuentre dentro de los números reales. Esta distribución también es llamada distribución de Gauss o gaussiana y se basa en dos parámetros, la media y la desviación estándar para las cuales se asumen valores de 1 y 0 respectivamente, esto gráficamente puede representarse por una campana simétrica (Campana de Gauss) para la función densidad de los datos de la serie.

La distribución normal, para el planteamiento de las pruebas de hipótesis, hace una comparación entre dos parámetros Z, los cuales son trazados en el intervalo de la campana y según la naturaleza de la hipótesis planteada ésta puede ser rechazada o aceptada. Los valores de

Z utilizados son calculados a partir del valor de la probabilidad analizada y buscada en la tabla ya creada para la distribución normal, que va de valores de $-\infty$ a $+\infty$ y el valor calculado de Z se determina de la siguiente manera:

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \text{ [Ec. 9]}$$

Con estos dos valores de Z se puede aceptar o rechazar la hipótesis nula planteada, al evaluar si el valor de Z que fue calculado queda dentro de la zona de rechazo establecido por el valor de Z dado por la tabla de distribución normal. (Noste, 2013, p. 51).

2.1.2.5. Principio de parsimonia

El concepto de parsimonia estadística consiste básicamente, en el criterio lógico que indica que al reducir en la manera de lo posible la solución a un problema, ésta será óptima. En estadística, este principio se respalda en el hecho de que, a mayor cantidad de variables introducidas en un modelo, se restan los grados de libertad disponibles y se aumenta el riesgo de generar ruido en el modelo, así como multicolinealidad si se desarrolla un modelo de regresión múltiple.

El principio de parsimonia también es conocido como de “Navaja de Ockham” nombre, debido a su creador, un fraile, filósofo y lógico

franciscano del siglo XIV, Guillermo Ockham, quien formuló la teoría que establece que, si para la explicación de un fenómeno se cuenta con varias hipótesis propuestas, lo más lógico sería escoger que quien lo explica mejor es la más sencilla de todas. (Barón, 2004, p. 98).

En estadística este principio es usado para simplificar modelos de regresión lineal múltiple, a manera que basados en la colinealidad de las variables independientes que forman el modelo pueda reducirse el número de ellas y crear un modelo matemático predictivo más simple, más fácil de validar y que de igual manera sea representativo de lo que se desea pronosticar. (Morales, 2012, p. 95).

2.1.3. Regresión logística binomial

La regresión logística binomial es un método estadístico correlacional utilizado para conocer las probabilidades de que se cumpla o no un evento (en la mayoría de los casos el padecimiento de alguna enfermedad o condición de salud), en función de un conjunto o conjuntos de variables regresoras cuantitativas o cualitativas.

La Regresión Logística Binaria (RLB) se usa cuando se desea conocer la relación entre una variable dependiente cualitativa dicotómica (dependencia) y una o más variables independientes o explicativas, que pueden ser cualitativas (variables sociodemográficas) y/o cuantitativas,

con el objetivo de obtener una estimación ajustada de la probabilidad de ocurrencia de un evento a partir de una o más variables independientes. Cuando se aplica a distintas variables independientes o con múltiples categorías, se llama Regresión Logística Binaria Multivariante (RLBM), cuya característica más interesante es la relación que tienen los coeficientes obtenidos, en la cuantificación del riesgo mediante el Odds Ratio (OR). (García, García, González y García. 2010, p. 366).

Para obtener las probabilidades de que se cumpla o no el evento correspondiente a la variable dicotómica, se utiliza la siguiente ecuación que corresponde a los logaritmos de las probabilidades binomiales en función de los coeficientes de cada variable regresora (categórica o numérica):

$$\text{logit}(p_i) = \ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 X_{1,i} + \dots + \beta_k X_{k,i} \quad [\text{Ec. 10}]$$

2.1.4. Validadores de la regresión logística

Para evaluar el ajuste y efectividad de la regresión logística es necesario validar algunos supuestos estadísticos posteriormente al proceso de análisis de datos.

2.1.4.1. Análisis de devianzas

Una técnica para medir el ajuste de modelos lineales generales de regresión es el análisis de devianzas. El análisis consiste en comparar los

logaritmos de verosimilitud del modelo de estudio y de un modelo saturado, en el que se asume que hay tantos parámetros como observaciones.

La devianza, denotada por $D(y, \mu)$, es una medida de bondad de ajuste que compara el modelo saturado con el modelo de interés. La idea es la misma que se usa en la prueba de LRT pero con los dos modelos anteriores. Es decir,

$$D(y, \mu) = -2[\ln(\text{modelo}) - \ln(\text{saturado})] = -2[\ln(\hat{\mu}; y) - \ln(\hat{\mu}; \mu)]$$

[Ec.11]

donde $\ln(\hat{\mu}; y)$ es el logaritmo de la verosimilitud del modelo de interés y $\ln(\hat{\mu}; \mu)$ es el logaritmo de la verosimilitud del modelo saturado. Un modelo saturado es aquel en el cual se asume que hay tantos parámetros como observaciones, es decir, $\hat{\mu}_i = y_i$. (Dobson & Barnett, 2008, p. 81)

2.1.4.2. Estadístico de chi cuadrado

Otra prueba de bondad de ajuste para modelos lineales generales es el estadístico de chi cuadrado, en la que se utilizan comparaciones de valores observados y ajustados del modelo de regresión, para cada combinación de niveles o factores de variables regresoras.

Otra medida de bondad de ajuste en MLG es el estadístico x^2 , definido como:

$$X^2 = \sum_i \frac{(y_i - \hat{\mu}_i)^2}{Var((y_i))} \text{ [Ec. 12]}$$

La distribución de X^2 se puede aproximar a una ji-cuadrado con grados de libertad igual al número de observaciones menos el número de parámetros en el modelo. Cuando el número de observaciones es suficiente grande, los estadísticos D(devianza) y X^2 son equivalentes. Algunos autores recomiendan el uso de X^2 en lugar de D ya que esta última se deja afectar más cuando las frecuencias esperadas son muy pequeñas. (Dobson & Barnett, 2008, p. 114)

2.1.4.3. Estadístico de Wald

Para validar el porcentaje de explicación que aporta cada variable en un modelo de regresión de forma individual, se utiliza el estadístico de Wald, que sigue una distribución normal.

Como en la regresión lineal, queremos saber no sólo cómo de bueno es el modelo ajustándose a los datos, sino también la contribución individual de cada una de las variables predictoras. Esta información la proporciona el *estadístico de Wald* (z-statistic) que sigue una distribución normal.

Al igual que el *estadístico t* en regresión lineal, el *estadístico z* nos dice si los coeficientes b_s para cada predictora son significativamente diferentes de cero. Si es distinto de cero asumimos que la variable predictora está haciendo una contribución significativa al modelo para predecir la respuesta (Y).

Su valor para un coeficiente concreto viene dado por el cociente entre el valor del coeficiente y su correspondiente error estándar. La obtención de significación indica que dicho coeficiente es diferente de 0 y merece la pena su conservación en el modelo.

Sin embargo, en modelos con errores estándar grandes, el estadístico de Wald puede proporcionar falsos resultados. Si el coeficiente (b) del modelo es grande, el error estándar tiende a inflarse y esto incrementa la probabilidad de rechazar un predictor cuando en realidad está haciendo una contribución al modelo.

El estadístico de Wald se determina mediante la relación entre la diferencia del valor estimado de b respecto al valor de contraste ($b=0$) y el error típico. (Rioja, Almazán, y Mañas. 2013, p. 20)

$$Wald = \frac{(\hat{b} - \hat{b}_0)^2}{(ET_{\hat{b}})^2} \text{ [Ec. 13]}$$

2.1.4.4. Odds Ratio

Los Odds Ratio son medidas que permiten comparar las probabilidades o riesgos de que suceda un evento entre diferentes grupos o niveles de una misma variable explicativa. Se utilizan frecuentemente en estudios de salud. Pueden ser calculados en diseños prospectivos, retrospectivos y transversales.

Odds Ratio (OR) es una medida de efecto comúnmente utilizada para comunicar los resultados de una investigación en salud. Matemáticamente un OR corresponde a un cociente entre dos odds, siendo un odds una forma alternativa de expresar la posibilidad de ocurrencia de un evento de interés o de presencia de una exposición. Desde un punto de vista metodológico, los OR pueden ser calculados en diseños prospectivos, retrospectivos y transversales, y bajo ciertas condiciones pueden reemplazar al Riesgo Relativo. En base a una serie de preguntas y ejemplos, el presente artículo explica los fundamentos matemáticos y metodológicos subyacentes al concepto de OR, con el objetivo de facilitar su interpretación a clínicos e investigadores. (Cerdeira, Vera y Rada. 2013, p. 1329).

En los modelos de regresión logística, los Odds Ratio desempeñan un papel importante como salidas y validadores. Los Odds Ratio se pueden calcular aplicando una función de logaritmo inverso a cada uno de los coeficientes beta (β), de la función de regresión obtenida. De esta forma se pueden comparar los riesgos de presentar o no una condición o evento de respuesta entre diferentes

niveles de cada variable explicativa, al mismo tiempo que se evalúa el modelo en función de los OR obtenidos. Al saber que el valor 1 para cada OR obtenido significa diferencia nula entre los riesgos de obtener el resultado pronosticado, se busca obtener valores alejados a 1 para identificar diferencias significativas entre los niveles de cada factor de regresión, y así, conocer, las variables significativas y los niveles que representen un riesgo estadístico.

Si nuestros datos se ajustan de manera satisfactoria a este modelo, tendremos la suerte de poder explicar la relación entre las variables independientes y la respuesta de una manera muy sencilla. Los coeficientes β_i expresan el logaritmo neperiano del odds ratio (OR) para cada factor de riesgo X_i . Por tanto, el OR se estima a partir de la fórmula:

$$OR = \text{antilog}(\beta_i) = e^{\beta_i} \text{ [Ec. 14]}$$

Para poder interpretar una OR, es necesario siempre tener en cuenta cuál es el factor o variable predictora que se estudia y cuál es el resultado o desenlace. El valor nulo para la OR es el 1, esto implica que las dos categorías comparadas son iguales. También recibe el nombre de «razón de odds» o «razón de ventajas». (Fiuza y Rodríguez, 2000)

2.1.4.5. Criterio de información AIC

El criterio de información de Akaike (AIC), utiliza la función de máxima verosimilitud y la cantidad de parámetros aplicados en un modelo de regresión, para obtener un estimador insesgado del ajuste de los datos. Entre menor es la magnitud relativa del criterio, se puede afirmar con mayor seguridad que el modelo desarrollado es efectivo al predecir y explicar la variable respuesta de acuerdo con la función de verosimilitud, y se mantiene bajo un nivel aceptable de complejidad en función de la cantidad de parámetros, respetando el principio de parsimonia.

El criterio de información Akaike propuesto en 1974 y utilizado como un estimador insesgado, precisa que el modelo de probabilidad de menor valor AIC es el que se selecciona como el mejor al que se le ajustan los datos. La función está dada por la maximización del logaritmo de la máxima verosimilitud denotado como $\ln(L)$ y K es el número de parámetros de la función de probabilidad (parámetros en el modelo):

$$AIC = 2K - 2\ln(L) \text{ [Ec. 15]}$$

La estructura del AIC está compuesta por la maximización del logaritmo de verosimilitud, es decir, como componente de la falta de ajuste del modelo y K como el número de parámetros estimados dentro del modelo como componente de penalidad. La penalidad es una medida de la complejidad o compensación por el sesgo debido a la falta de ajuste

cuando los estimadores de máxima verosimilitud son empleados.
(Calderón, Navarrete y Diaz, 2015)

2.1.4.6. Prueba de Hosmer Lemeshow

La prueba de Hosmer Lemeshow es otra prueba de bondad de ajuste para modelos de regresión logística binomial, que consiste en utilizar la distribución chi cuadrado para calcular las probabilidades relacionadas al valor respuesta de la variable dependiente para cada una de las observaciones de un conjunto de datos. Con las probabilidades calculadas para cada observación, se puede obtener una significancia p al comparar las frecuencias observadas con las esperadas. Si el valor de significancia es mayor a α , se puede decir que el ajuste del modelo es apropiado.

La prueba de Hosmer Lemeshow evalúa la bondad del modelo construyendo una tabla de contingencia, divide la muestra en aproximadamente 10 grupos iguales a partir de las probabilidades estimadas, para comparar las frecuencias observadas con las esperadas en cada uno de estos grupos a través de la prueba χ^2 con $j - 2$ grados de libertad, en donde j es el número de grupos formados.

Se calculan los deciles de las probabilidades estimadas \hat{P}_i , divididos en 10 grupos dados por:

$$A_j = \{i \in \{1, \dots, n\} / \hat{P}_i \in [D_{j-1}, D_j]\}, j = 1, \dots, 10 \text{ [Ec. 16]}$$

$$\text{Donde } D_0 = 0 \text{ y } D_{10} = 1$$

Luego, la probabilidad estimada para cada grupo está dada por:

$$\bar{P}_j = \frac{1}{n_j} \sum_{i \in A_j} \hat{P}_i ; j = 1, \dots, 10 \text{ [Ec. 17]}$$

Donde n_j indica el número de casos del grupo A_j . (Hosmer et al, 1997, p. 965)

Cada probabilidad estimada por grupo se contrasta con los valores observados relacionados al valor 1 de la variable respuesta, mediante la prueba chi cuadrado. Si el valor de significancia es mayor al alfa definido, se acepta la hipótesis nula de la bondad de ajuste del modelo.

2.1.5. Plan de muestreo

Para los estudios de regresión múltiple y logística, no existe un plan de muestreo definido. Sin embargo, al considerar que la regresión logística binomial tiene como objetivo determinar si los miembros de una muestra cumplen o no con una condición o evento, es aceptable la utilización de muestreo proporcional, el cual se basa en el principio de cálculo de número de individuos necesario para determinar qué proporción de una población cumple o no con una característica binomial o dicotómica. Cuando la proporción o probabilidad de que se cumpla o no con la condición o evento es desconocida, se utiliza el valor $p = 0.5$, para incluir el mayor número de muestras.

Las fórmulas para calcular el tamaño de la muestra cuando el objetivo radica en estimar la proporción poblacional. Si el tamaño de la población es conocido (población finita):

$$n = \frac{k^2 pqN}{e^2(N-1) + k^2 pq} \text{ [Ec. 18]}$$

Si el tamaño de la población es desconocido (población infinita):

$$n = \frac{k^2 pq}{e^2} \text{ [Ec. 19]}$$

Nomenclatura: n = Tamaño de la muestra. N = Total de elementos que integran la población. k^2 = Zeta crítico: valor determinado por el nivel

de confianza adoptado, elevado al cuadrado. e = Error muestral: falla que se produce al extraer la muestra de la población. Generalmente, oscila entre 1% y 5%. p = Proporción de elementos que presentan una determinada característica a ser investigada. Una proporción es la relación de una cantidad con respecto a otra mayor. Por ejemplo, en un grupo de 100 estudiantes hay 75 mujeres y 25 hombres. La fórmula es $p = A/N$. Entonces la proporción de mujeres es $75/100 = 0,75$ y la proporción de hombres es $25/100 = 0,25$. q = Proporción de elementos que no presentan la característica que se investiga. (Arias, 2012, p. 89)

2.2. Indicadores nutricionales infantiles y variables RV

Como parte de los fundamentos técnicos del estudio, se consideran conceptos relacionados a los métodos de clasificación de condiciones de nutrición y malnutrición infantil. Asimismo, se consideran conceptos relacionados a las condiciones de vivienda en términos de infraestructura e higiene.

2.2.1. Medición nutricional infantil

La medición nutricional infantil es un aspecto importante para poder dar un seguimiento y clasificación de condiciones nutricionales correctas a nivel individual y colectivo. La medición permite definir criterios de clasificación y aplicar medidas de mejora en función de las condiciones nutricionales en que se encuentren los individuos.

2.2.1.1. Malnutrición infantil

La malnutrición infantil es una condición de salud y socioeconómica que está presente en muchos países en vías de desarrollo y en países desarrollados, que atraviesan crisis económicas. Esta condición puede aumentar los riesgos de padecer enfermedades respiratorias e intestinales en niños, lo que, a su vez, provoca efectos a largo plazo e incluso la muerte.

La malnutrición infantil es un problema a nivel mundial, tanto por defecto como por exceso. La desnutrición persiste en los países emergentes y comienza a evidenciarse en algunos sectores de los países desarrollados debido a la crisis económica.

Los indicadores antropométricos reflejan el estado nutricional del individuo a través del tamaño, composición y proporcionalidad corporal. El contraste de las dimensiones individuales con los valores de referencia que corresponden a un menor sano de la misma edad y sexo permite establecer un diagnóstico del tipo y grado de malnutrición. (Díez & Marrodán, 2018, p. 27)

La condición de desnutrición o malnutrición infantil predispone a la presencia de enfermedades diarreicas y respiratorias en los niños, las cuales surgen posteriormente como las primeras causas de morbilidad y

mortalidad, cuando en realidad constituyen fenómenos agregados al problema de la desnutrición. (Tovar, Navarro y Fernández. 1997, p. 47)

2.2.1.2. Clasificación del estado nutricional infantil

La clasificación de Waterlow es de las más utilizadas para los diagnósticos nutricionales infantiles por su practicidad y efectividad. Utiliza las mediciones de índices antropométricos para definir si un individuo está en condiciones de desnutrición crónica o aguda.

La clasificación de Waterlow permite determinar la cronología y la intensidad de la desnutrición. Para la realización de esta evaluación se necesitan dos indicadores:

$$\text{Porcentaje de } \frac{\text{peso}}{\text{estatura}} \left(\frac{P}{E} \right) = \frac{\text{peso real}}{\text{peso que debería tener para la estatura}} \times 100$$

[Ec. 20]

$$\text{Porcentaje de } \frac{\text{estatura}}{\text{edad}} \left(\frac{T}{E} \right) = \frac{\text{estatura real}}{\text{estatura que debería tener para la edad}} \times 100$$

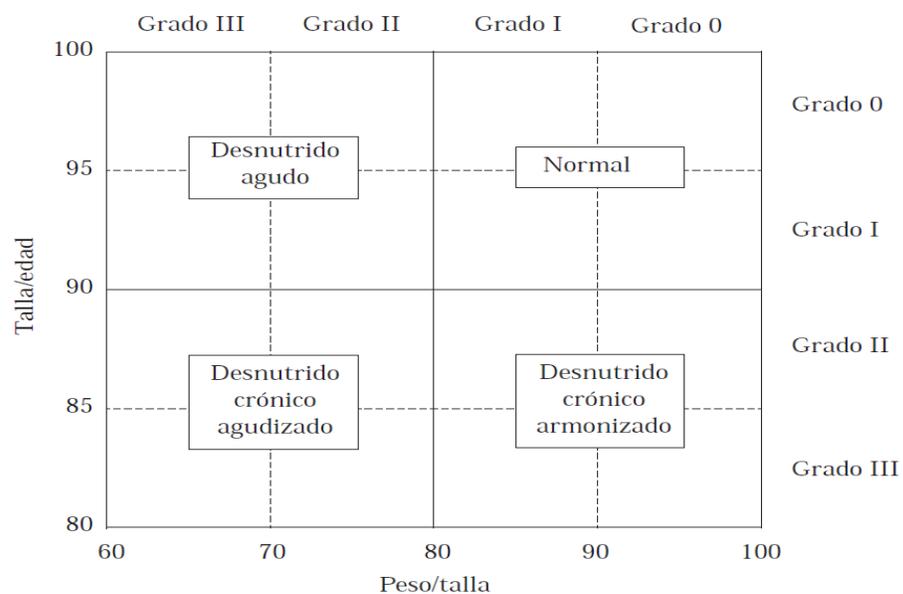
[Ec. 21]

Con el fin de determinar el peso para la talla y la talla para la edad, es necesario tomar el valor que corresponde al percentil 50 en las gráficas de crecimiento.

Una vez que se han obtenido los porcentajes, se ubican según el criterio siguiente: nutrición normal (cuando el peso para la talla y la talla para la edad se encuentran dentro de valores adecuados para la edad), desnutrición aguda (peso para la talla bajo y talla para la edad normal), desnutrición crónica recuperada o en homeorresis (talla para la edad alterada y peso para la talla normal), desnutrición crónica agudizada (talla para la estatura alterada y peso para la talla baja). Por densidad la clasificación será: Grado I (menos del 90 %), grado II (entre el 80 y 89 %) y grado III (menos del 79 %). (Márquez-González et al, 2012, p. 66)

Figura 1.

Clasificación nutricional por Waterlow



Nota. Obtenido de H.M. (2012). *Clasificación de Waterlow*. (medigraphic.com/pdfs/residente/rr-2012/rr122d.pdf), consultado el 21 de septiembre de 2021. De dominio público.

De manera similar al método de Waterlow, la clasificación por desviaciones está establecida por la OMS.

2.2.1.3. Utilización de índices antropométricos

Los valores de antropometría constituyen una importante fuente de información para evaluar la condición nutricional de personas a cualquier edad, especialmente en niños. Asimismo, permiten la evaluación y seguimiento en el tiempo. Esto se debe a que el desarrollo del peso y de la talla de las personas es dependiente de la edad que poseen.

El crecimiento y desarrollo de los niños es el resultado de las características genéticas y las condiciones ambientales a las que están expuestos, en un proceso que no es uniforme, en el cual hay etapas de mayor velocidad en donde los niños son más susceptibles a presentar problemas nutricionales como la desnutrición, el sobrepeso y la obesidad. Por ello, es importante utilizar los índices antropométricos adecuados a las características del crecimiento de los niños. Convencionalmente se han utilizado los índices P/E, T/E y P/T en niños menores de 5 años. Actualmente se recomienda también el uso del IMC. Un índice nutricional debe tener la capacidad de identificar a los sujetos bien nutridos de los que presentan alguna alteración nutricional por déficit o por exceso (desnutrición y sobrepeso y obesidad); por lo tanto, se requiere que se

utilicen los índices nutricionales apropiados para la edad de la población a evaluar. (Ochoa, García, Flores, García-Miranda, y Solís. 2017, p. 823).

La talla para la edad (T/E) informa sobre la desnutrición crónica o retraso del crecimiento, el peso para la edad (P/E) acerca de la desnutrición global y tanto el peso para la talla (P/T) como el perímetro del brazo identifican la desnutrición aguda. (Díez & Marrodán, 2018, p. 27).

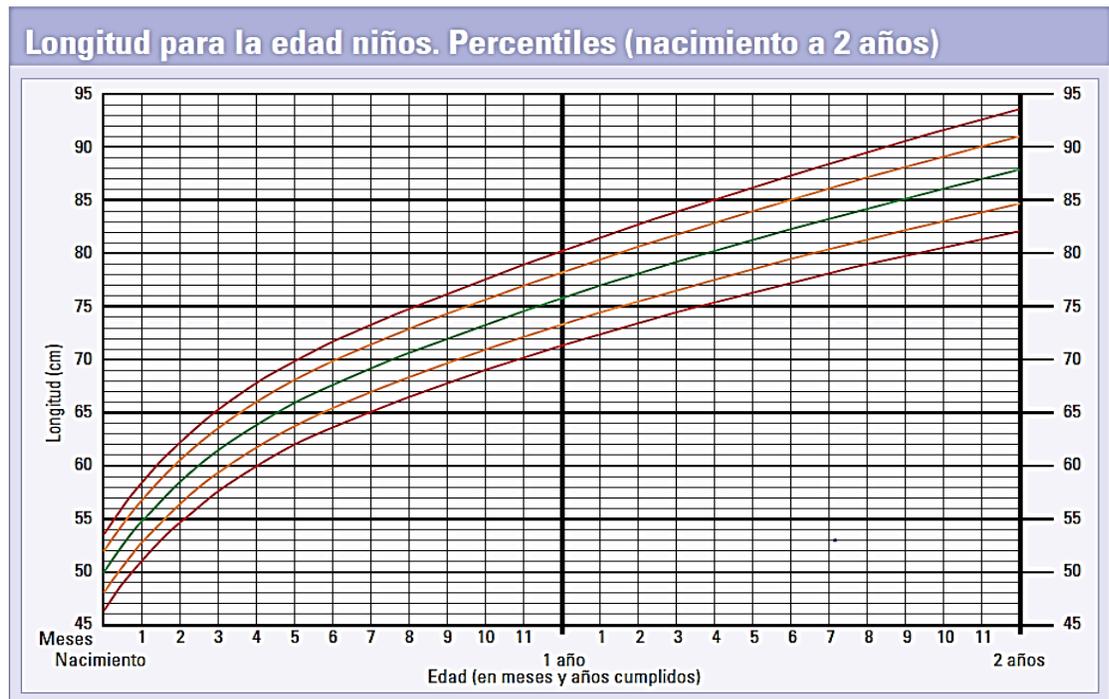
- Gráfica de talla para la edad (T/E)

La talla respecto a la edad es usada para identificar niños con bajo crecimiento o crecimiento lento respecto al percentil correspondiente a la media poblacional. Es importante el seguimiento de este indicador en los primeros 5 años.

Refleja el crecimiento lineal alcanzado en relación con la edad cronológica. Nos indica si hubo problemas relacionados con la nutrición en el pasado. Un valor fuera de los límites indica que el niño no tiene la longitud o talla que debería tener para su edad y que hay un problema nutricional. (INCAP, 2012, p. 10)

Figura 2

Longitud para la edad niños



Nota. Obtenido de OMS (2017). *Longitud para la edad niños. Percentiles (nacimiento a 2 años)*. (https://www.aepap.org/sites/default/files/curvas_oms.pdf), consultado el 21 de septiembre de 2021. De dominio público.

- Gráfica de peso para la edad (P/E)

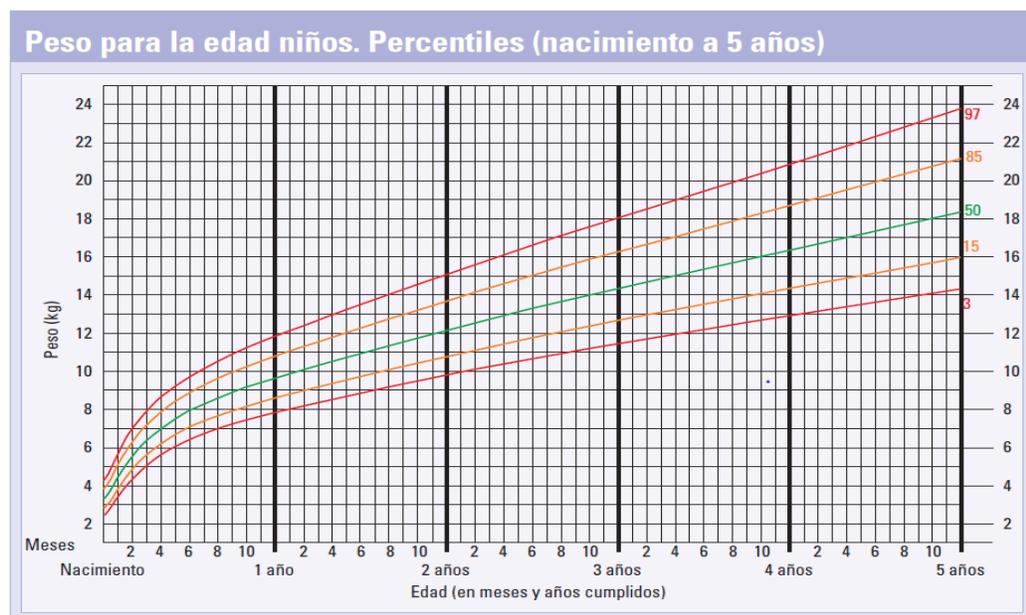
El indicador de peso para la edad permite identificar casos de posibles problemas alimenticios durante el primer año. Asimismo, sirve para indicar si un niño tiene condición de peso bajo o peso bajo excesivo.

Refleja la masa corporal (peso) alcanzada para la edad cronológica. Es un indicador muy útil para monitorear el crecimiento de un niño, sobre todo en el primer año. Un valor fuera de los límites indica que el niño no tiene el peso que debería tener para su edad y que hay un problema nutricional.

Si el resultado es bajo, indica que existe déficit de peso de acuerdo con el que debería tener para su edad. Si el resultado es alto, es un indicador de exceso de peso de acuerdo con el que debería tener para su edad. (INCAP, 2012, p. 10)

Figura 3.

Peso para la edad niños



Nota. Obtenido de OMS (2017). *Peso para la edad niños. Percentiles (nacimiento a 5 años)*. (https://www.aepap.org/sites/default/files/curvas_oms.pdf), consultado el 21 de septiembre de 2021. De dominio público.

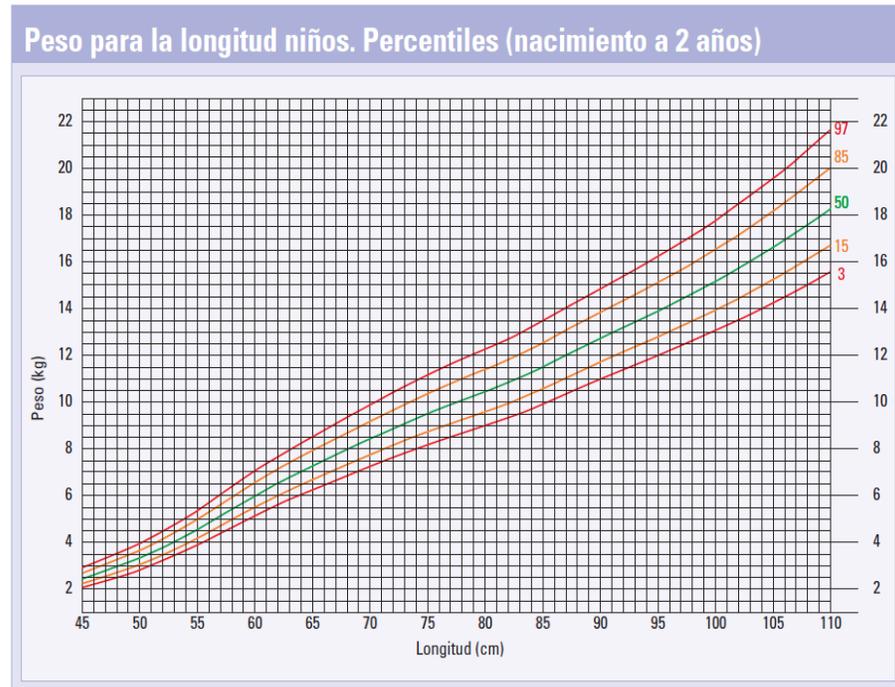
- Gráfica de peso para la talla (P/T)

El peso para la talla sirve como una medida del crecimiento y de la salud de los niños, al asumir hasta cierto grado que la estatura presenta un comportamiento normal. Usualmente se utiliza para identificar problemas de emaciación o adelgazamiento patológico.

Refleja el peso relativo para la longitud o talla del niño. Nos indica el estado nutricional actual del niño. Un valor fuera de los límites indica que el niño no tiene el peso que debería tener para la longitud o talla que ha alcanzado a la fecha de la medición. Si el resultado es bajo, indica que existe un déficit en el peso que tiene para la longitud o talla alcanzada. Si el resultado es alto, es un indicador de exceso y significa que el niño tiene más peso del que debería tener para la estatura alcanzada a la fecha de la medición. (INCAP, 2012, p. 10).

Figura 4.

Peso para la talla niños



Nota. Obtenido de OMS (2017). *Peso para la talla niños. Percentiles (nacimiento a 2 años)*. (https://www.aepap.org/sites/default/files/curvas_oms.pdf), consultado el 21 de septiembre de 2021. De dominio público.

Las gráficas de crecimiento infantil de la Organización Mundial de la Salud –OMS– son una herramienta útil para registrar y evaluar el crecimiento físico de los niños a través de los años. Esto tiene el propósito de conocer el ritmo de crecimiento para ayudarlo a desarrollar su máximo potencial en talla, peso y, a la vez, puedan detectarse y corregir oportunamente posibles alteraciones.

Estas gráficas de crecimiento tienen trazados varios canales (o curvas) que representan una serie de valores de peso y talla que describen el crecimiento promedio de niños. Estos canales o curvas se dividen la línea mediana, que representa el valor ideal en el que debería estar ubicada la medición de un niño a la hora de marcar el punto sobre la gráfica y en las desviaciones estándar (DE), que indican la distancia que se aleja la medición de un niño de la mediana o valor ideal. (INCAP, 2012, p. 11)

- Índice de masa corporal (IMC)

El índice de masa corporal ha sido utilizado principalmente en adultos para identificar condiciones tanto de sobrepeso como de obesidad. Sin embargo, también se aplica para evaluar condiciones nutricionales en niños y adolescentes en cuanto a corpulencia, que, en niños con condiciones de desnutrición, es considerada como delgadez o delgadez extrema.

El IMC desarrollado por Quetelet en 1869 se ha utilizado principalmente en adultos y, en fechas recientes, para evaluar el estado nutricional de niños y adolescentes. Para su cálculo se utiliza la siguiente fórmula:

$$IMC = \frac{peso[kg]}{talla^2[m]} \text{ [Ec. 22]}$$

El IMC es un indicador simple de la relación entre el peso y la talla que se utiliza frecuentemente para identificar el sobrepeso y la obesidad. Por su simpleza, bajo costo y adecuada correlación con la grasa corporal total, la Organización Mundial de la Salud (OMS) lo recomienda también para evaluar antropométricamente el estado nutricional de la población menor de 20 años.

Sin embargo, no se debe olvidar que es un índice global de corpulencia que no permite diferenciar la masa grasa de la masa magra, y por tanto, no es una medida precisa de adiposidad a nivel individual. Actualmente, es uno de los indicadores más utilizados para caracterizar el estado nutricional, no solo en adultos sino también en niños y adolescentes, y ha sido empleado tanto en estudios epidemiológicos como clínicos. (Ochoa et al, 2017, p. 821)

2.2.2. Características de vivienda e higiene (variables RV)

De acuerdo con varias entidades internacionales (como la OMS), las características de higiene y vivienda desempeñan un papel determinante a mediano y largo plazo en las condiciones nutricionales de todos los integrantes de la familia, especialmente los niños.

2.2.2.1. Definición y morfología de vivienda

La vivienda representa el espacio físico que alberga a los integrantes de una familia, con los componentes necesarios para realizar sus actividades tanto a nivel personal como colectivo.

Vivienda, es la edificación o construcción material con fines de refugio y habitación para una familia o personas. Siempre ha sido una necesidad primordial que las sociedades procuran satisfacer para que los individuos desarrollen sus aspectos físicos, sociales, emocionales y espirituales.

La morfología de la vivienda en Guatemala, específicamente en el área rural, se basa en la función; es la respuesta a las necesidades del lugar y a la capacidad del propietario para construir. Normalmente se establece un patrón de morfología por la distribución de la planta de la vivienda y de los elementos verticales tales como las ventanas, puertas, techo, alturas, y detalles simbólicos o estéticos. (Mejía, Linares y Aguilar. 2006, p. 26)

2.2.2.2. Sistemas y métodos constructivos

El sistema constructivo es una forma de clasificación estructural basada en la disposición o combinación de materiales utilizados para construir las diferentes partes de una vivienda, como el techo, piso, paredes, entre otros.

El sistema constructivo es la forma y tecnología para construir, regida también por los diferentes materiales para trabajar.

Los sistemas y métodos constructivos que se describen a continuación son los más aplicados en nuestro país, tanto en el área urbana como rural, con la salvedad que existen otras tecnologías, en etapas de experimentación.

Los sistemas constructivos utilizados en Guatemala son:

- Construcción de tierra y adobe

Este método constructivo es de los más comunes en el área rural por la disponibilidad de los materiales que se requieren. Es importante considerar el factor de humedad en la tierra para que la construcción sea resistente.

Este método constructivo requiere básicamente tierra y agua. Es aplicado en gran escala en el área rural. Actualmente se han continuado varios estudios para mejorar las cualidades de la tierra; la tierra es el único material accesible para edificaciones en muchas regiones y países. Uno de los factores más importantes para trabajar con tierra es el contenido de humedad de ésta; permite la cohesividad de sus partículas.

Así también, la mezcla de sus componentes primarios, es decir otras tierras. Por otro lado, si la tierra está seca no permitirá el enlace cohesivo de sus partículas que provienen de la integración de la arcilla y la humedad. Las capas secas de material no consolidado también producirán planos estructurales débiles dentro de la estructura del muro.

- Construcción de madera

La madera es muy común en las viviendas rurales con grandes áreas boscosas y su construcción es simple y liviana.

La madera es un material apropiado para los más variados aspectos de la construcción. Sus características constructivas son su fácil trabajo, escaso peso, ventajosas cualidades estáticas, cualidades artísticas, entre otras. A través del tiempo se han desarrollado químicos para fortalecer la madera contra la putrefacción, los animales nocivos y el fuego.

- Construcción mixta

Las construcciones mixtas son muy frecuentes en las áreas rurales de Guatemala y muchos otros países de Latinoamérica. Las paredes regularmente de blocks o ladrillos de barro complementado con cemento, arena, pedrín y lámina.

Es el sistema tradicional de nuestro país y de los países en vías de desarrollo. Aunque su aplicación como sistema no tiene muchas décadas. Con este sistema constructivo da como resultado edificaciones sólidas, seguras y su comportamiento estructural puede ser calculado con precisión. Las construcciones con este sistema pueden tener espacios amplios sin apoyos intermedios gracias a las facultades del concreto. Además, este tipo de construcción no exige mucho mantenimiento.

Las paredes son de bloques colocados sucesivamente uno encima de otro y uno al lado del otro. Los principales materiales son cemento,

arena, pedrín, block, acero. Una variedad de este sistema es el ladrillo de barro cocido, que en realidad tiene características similares a las del block.

- **Construcción prefabricada**

Los sistemas prefabricados han incrementado su utilización en los últimos años, por su fácil construcción e instalación. Sus componentes pueden variar, pero generalmente ofrecen buena durabilidad y condiciones cómodas para los habitantes de la vivienda.

Estos sistemas son prácticamente modernos y su característica principal es que las partes de la edificación ya han sido construidas en una fábrica de producción en serie, lo cual suele reducir tiempos de ejecución, a la vez que facilita el trabajo. Dentro de estos sistemas constructivos tenemos los prefabricados en madera, concreto, metal, polietileno y poliuretano (paneles), madera enchapada con variedad de productos, asbesto de cemento, como los más conocidos. (Mejía et al, 2006, p. 29).

2.2.2.3. Abastecimiento de agua salubre

El agua potable o salubre, de acuerdo con la OMS, constituye un factor de suma importancia para las condiciones de salud e higiene en la vivienda. Tanto la fuente como el método de purificación inciden en la calidad del agua utilizada.

El abastecimiento de suficiente agua salubre y potable contribuye a prevenir la propagación de enfermedades gastrointestinales, propicia la higiene doméstica y personal y mejora el nivel de vida. Como el agua es esencial para vivir, se la debe proteger contra la contaminación biológica por microorganismos nocivos como la *Shigella*, la salmonela, la *Escheridia coli interpatógena*, ciertos virus eitróricos y diversos protozoos y helmintos parasitarios.

Las enfermedades diarreicas transmitidas por el agua afectan especialmente a los niños pequeños y, en algunos países en desarrollo, pueden causar hasta la tercera parte de las defunciones de niños menores de cinco años. La gravedad de esas enfermedades aumenta notablemente cuando la exposición se combina con los efectos de la malnutrición. Para una higiene personal y doméstica adecuada se necesita agua en cantidad razonable, la cual, suministrada en condiciones idóneas, promueve esos usos, así como la productividad de la familia y la preparación higiénica de los alimentos. (OMS, 1990, p. 3).

2.2.2.4. Eliminación higiénica de excretas

La exposición de material fecal puede ser causa de diversas enfermedades e infecciones que, a largo plazo, pueden provocar condiciones de desnutrición infantil. La presencia de heces fecales puede ser por exposición de

las diversas fuentes de agua o en cultivos que se usen para el sustento de las familias.

Una de las vías principales de contaminación biológica del agua, los alimentos y el suelo son las heces humanas. La contaminación puede producirse cerca de las casas, por ejemplo, cuando la gente defeca en el suelo o en terrenos dedicados al cultivo de productos alimenticios o cuando las letrinas están mal situadas en relación con los pozos, construidas en suelo carente del necesario drenaje o mal conservadas.

El rebose de las letrinas enloda los alrededores con la consiguiente exposición directa a helmintos y protozoos parasitarios y a otros organismos patógenos y fomenta la abundancia de moscas. La exposición puede ser menos directa, por ejemplo, cuando se introducen en las fuentes de agua, y luego en la cadena alimentaria, excretas no tratadas que transmiten organismos patógenos. (OMS, 1990, p. 4).

2.2.2.5. Eliminación de desechos sólidos

La forma de eliminar los desechos sólidos es determinante para las condiciones de higiene y de salud de la vivienda, porque si no se hace de forma apropiada. Por ejemplo, si se queman en un área cercana y se generan emanaciones de humo, se pueden provocar condiciones que aumentan el riesgo de contagio de enfermedades respiratorias y gastrointestinales, lo que, a su vez, perjudica las condiciones nutricionales de los integrantes de la familia.

La eliminación adecuada e higiénica de los desechos sólidos domésticos reduce los riesgos para la salud y crea un entorno más agradable para la vida; la utilización de métodos apropiados de almacenamiento y eliminación es desfavorable a los insectos y roedores vectores de enfermedades y protege a las personas contra las sustancias venenosas y los objetos que pueden causar traumatismos accidentales.

El inadecuado almacenamiento, recogida y eliminación de los desechos sólidos, fenómeno sobre todo urbano, puede dar lugar a una serie de riesgos para la salud, en particular la difusión de enfermedades gastrointestinales y parasitarias, sobre todo cuando hay excrementos humanos mezclados con otros desechos orgánicos. (OMS, 1990, p. 5).

2.2.2.6. Salvaguardias estructurales

La ubicación, estructura y amueblamiento de las viviendas son factores importantes para las condiciones de salud y bienestar en general, para las viviendas familiares.

La ubicación, la estructura y el amueblamiento adecuados de la vivienda protegen la salud, promueven la seguridad y reducen los riesgos. El tipo de vivienda disponible depende de las condiciones climáticas y económicas y de las preferencias culturales.

Al respetar las limitaciones de ese tipo, los proyectos, los materiales y las técnicas de construcción deben producir estructuras duraderas, que proporcionan un alojamiento seguro, seco y confortable y protejan a los moradores contra sabandijas, temperaturas extremas y riesgos naturales recurrentes (terremotos, huracanes, vientos). (OMS, 1990, p. 8).

3. PRESENTACIÓN DE RESULTADOS

De acuerdo con los objetivos propuestos, se presentan los resultados obtenidos con relación al modelo de regresión logística desarrollado para los factores de desnutrición infantil en el suroccidente de Guatemala.

3.1. Objetivo 1. Determinación de la medida en la que se relacionan estadísticamente la presencia de desnutrición infantil con las características físicas y energéticas de los hogares, para mejorar las propuestas de desarrollo social mediante el uso de análisis de devianza.

A continuación, se presentan los resultados obtenidos correspondientes al objetivo 1, que consiste en el cálculo de correlación entre las variables predictoras y la variable respuesta (presencia de desnutrición), de acuerdo con el análisis de devianza elaborado a partir de la distribución chi cuadrado.

3.1.1. Cálculo de diagnóstico

Para realizar el cálculo de diagnóstico de desnutrición con la información recopilada, se utilizó el método de diagnóstico de Waterlow y el método de desviaciones estándar implementado por la OMS. En la tabla II muestra un extracto de la información antropométrica de cada niño, utilizada para realizar el cálculo de diagnóstico con ambos métodos.

Tabla 10.

Extracto información antropométrica

Id	género	edad meses	edad días	peso kg	talla cms
320	M	4	125	5.9	58.0
113	H	23	705	10.8	82.0
407	M	19	581	8.4	79.0
408	M	10	288	6.4	62.0
411	H	22	592	9.6	79.0

Nota. Elaboración propia, realizado con Microsoft Excel.

3.1.1.1. Diagnóstico de Waterlow

Para realizar el cálculo de diagnóstico de Waterlow, se utilizaron las mediciones antropométricas presentadas en la tabla II, en conjunto con las tablas de talla para la edad (días) y peso para la talla en percentiles, publicadas por la OMS, para niños y niñas entre 0 y 24 meses. En las tablas III y IV se presentan extractos de los estándares publicados por la OMS para el seguimiento nutricional de talla para la edad. La edad en días de cada niño se calculó a partir de la fecha de nacimiento recopilada en contraste con la fecha de toma de datos antropométricos.

Tabla 11.*Extracto percentiles talla para la edad en niñas OMS*

Edad (días)	Media	Desviación estándar	P25	P50	P75	P90	P99
0	49.1477	0.0379	47.891	49.148	50.404	51.535	53.481
1	49.3166	0.03783	48.058	49.317	50.575	51.708	53.657
2	49.4854	0.03776	48.225	49.485	50.746	51.88	53.832
3	49.6543	0.0377	48.392	49.654	50.917	52.053	54.009
4	49.8232	0.03763	48.559	49.823	51.088	52.226	54.185
5	49.9921	0.03756	48.726	49.992	51.259	52.398	54.36
6	50.1609	0.03749	48.893	50.161	51.429	52.571	54.536
7	50.3298	0.03742	49.06	50.33	51.6	52.743	54.711
8	50.4987	0.03735	49.227	50.499	51.771	52.916	54.886
9	50.6676	0.03728	49.394	50.668	51.942	53.088	55.062
10	50.8365	0.03722	49.56	50.837	52.113	53.261	55.238
11	51.0053	0.03715	49.727	51.005	52.283	53.434	55.413
12	51.1742	0.03708	49.894	51.174	52.454	53.606	55.589
13	51.3431	0.03701	50.061	51.343	52.625	53.778	55.764
14	51.512	0.03694	50.229	51.512	52.795	53.951	55.939

Nota. Obtenido de OMS (2017). *Extracto percentiles talla para la edad en niñas.* (www.who.int/tools/child-growth-standards/standards/length-height-for-age), consultado el 21 de septiembre de 2021.

Tabla 12.*Extracto percentiles talla para la edad en niños OMS*

Edad (días)	Media	Desviación estándar	P25	P50	P75	P90	P99
0	49.8842	0.03795	48.607	49.884	51.161	52.31	54.288
1	50.0601	0.03785	48.782	50.06	51.338	52.488	54.468
2	50.2359	0.03775	48.957	50.236	51.515	52.666	54.648
3	50.4118	0.03764	49.132	50.412	51.692	52.844	54.826
4	50.5876	0.03754	49.307	50.588	51.868	53.021	55.005
5	50.7635	0.03744	49.482	50.764	52.045	53.199	55.185
6	50.9393	0.03734	49.656	50.939	52.222	53.377	55.364
7	51.1152	0.03723	49.832	51.115	52.399	53.554	55.542
8	51.291	0.03713	50.006	51.291	52.576	53.732	55.721
9	51.4669	0.03703	50.181	51.467	52.752	53.909	55.9
10	51.6427	0.03693	50.356	51.643	52.929	54.087	56.079
11	51.8186	0.03682	50.532	51.819	53.106	54.264	56.257
12	51.9944	0.03672	50.707	51.994	53.282	54.441	56.436
13	52.1702	0.03662	50.882	52.17	53.459	54.619	56.615
14	52.3461	0.03652	51.057	52.346	53.636	54.796	56.793
15	52.4978	0.03645	51.207	52.498	53.788	54.95	56.949
16	52.6488	0.03639	51.357	52.649	53.941	55.104	57.106
17	52.799	0.03633	51.505	52.799	54.093	55.257	57.261

Nota. Obtenido de OMS (2017). *Extracto percentiles talla para la edad en niños.* (www.who.int/tools/child-growth-standards/standards/length-height-for-age), consultado el 21 de septiembre de 2021. De dominio público.

En las tablas 13 y 14 se presentan extractos de los estándares publicados por la OMS para el seguimiento nutricional de peso para la talla. Las columnas de las tablas indican el percentil para cada medición de talla, entendiendo que la media equivale estadísticamente al percentil 50.

Tabla 13.

Extracto percentiles peso para la talla en niños OMS

Cm	3rd	15th	Media	85th	97th
45.0	2.1	2.2	2.4	2.7	2.9
45.5	2.1	2.3	2.5	2.8	3.0
46.0	2.2	2.4	2.6	2.9	3.1
46.5	2.3	2.5	2.7	3.0	3.2
47.0	2.4	2.5	2.8	3.1	3.3
47.5	2.4	2.6	2.9	3.1	3.4
48.0	2.5	2.7	2.9	3.2	3.5
48.5	2.6	2.8	3.0	3.3	3.6
49.0	2.7	2.9	3.1	3.4	3.7
49.5	2.7	2.9	3.2	3.5	3.8
50.0	2.8	3.0	3.3	3.7	4.0
50.5	2.9	3.1	3.4	3.8	4.1
51.0	3.0	3.2	3.5	3.9	4.2
51.5	3.1	3.3	3.6	4.0	4.3

Nota. Obtenido de OMS (2017). *Extracto percentiles peso para la talla en niños.* (<https://www.who.int/tools/child-growth-standards/standards/weight-for-age>), consultado el 21 de septiembre de 2021. De dominio público.

Tabla 14.*Extracto percentiles peso para la talla en niñas OMS*

cm	3rd	15th	Median	85th	97th
45.0	2.1	2.2	2.5	2.7	2.9
45.5	2.2	2.3	2.5	2.8	3.0
46.0	2.2	2.4	2.6	2.9	3.1
46.5	2.3	2.5	2.7	3.0	3.2
47.0	2.4	2.6	2.8	3.1	3.3
47.5	2.4	2.6	2.9	3.2	3.4
48.0	2.5	2.7	3.0	3.3	3.5
48.5	2.6	2.8	3.1	3.4	3.7
49.0	2.7	2.9	3.2	3.5	3.8
49.5	2.8	3.0	3.3	3.6	3.9
50.0	2.8	3.1	3.4	3.7	4.0
50.5	2.9	3.2	3.5	3.8	4.1
51.0	3.0	3.2	3.6	3.9	4.3
51.5	3.1	3.4	3.7	4.0	4.4

Nota. Obtenido de OMS (2017). *Extracto percentiles peso para la talla en niñas.* (<https://www.who.int/tools/child-growth-standards/standards/weight-for-age>), consultado el 21 de septiembre de 2021. De dominio público.

Utilizando las ecuaciones 17 y 18, se obtuvieron los porcentajes de P/T y T/E relativos de cada niño del estudio, en comparación con las medias de las tablas anteriores. De acuerdo con la clasificación presentada en la figura 1 para el diagnóstico de Waterlow, se obtuvieron las frecuencias de casos para cada nivel de clasificación, siendo estos normal (sin presencia de desnutrición), DCAR (desnutrición crónica armonizada), DA (desnutrición aguda), y DCAG (desnutrición crónica agudizada).

Tabla 15.*Clasificación de Waterlow para los casos estudiados*

Clasificación	Cuenta de casos	Porcentaje relativo
Normal	1481	76.74%
DCAR	433	22.44%
DA	10	0.52%
DCAG	6	0.31%
Total general	1930	100%

Nota. Elaboración propia, realizado con Microsoft Excel.

3.1.1.2. Diagnóstico según desviaciones estándar de la OMS

Para realizar el cálculo de diagnóstico de mediante el método de desviaciones estándar, se utilizaron las mediciones antropométricas presentadas en la tabla II, en conjunto con las tablas de talla para la edad (días) y peso para la talla en desviaciones estándar, publicadas por la OMS, para niños y niñas entre 0 y 24 meses. En las tablas XVI y XVII se presentan extractos de los estándares publicados por la OMS para el seguimiento nutricional de talla para la edad en desviaciones estándar (talla en centímetros).

Tabla 16.*Desviaciones estándar de talla para la edad en niños según OMS*

Day	SD4neg	SD3neg	SD2neg	SD1neg	SD0	SD1	SD2	SD3	SD4
0	42.312	44.205	46.098	47.991	49.884	51.777	53.67	55.564	57.457
1	42.481	44.376	46.271	48.165	50.06	51.955	53.85	55.744	57.639
2	42.65	44.547	46.443	48.339	50.236	52.132	54.029	55.925	57.822
3	42.822	44.719	46.617	48.514	50.412	52.309	54.207	56.104	58.002
4	42.991	44.89	46.789	48.689	50.588	52.487	54.386	56.285	58.184
5	43.161	45.062	46.962	48.863	50.764	52.664	54.565	56.465	58.366
6	43.331	45.233	47.135	49.037	50.939	52.841	54.743	56.646	58.548
7	43.503	45.406	47.309	49.212	51.115	53.018	54.921	56.824	58.727
8	43.673	45.578	47.482	49.387	51.291	53.195	55.1	57.004	58.909
9	43.844	45.749	47.655	49.561	51.467	53.373	55.279	57.184	59.09

Nota. Obtenido de OMS (2017). *Extracto desviaciones estándar talla para la edad en niños.* (www.who.int/tools/child-growth-standards/standards/length-height-for-age), consultado el 21 de septiembre de 2021. De dominio público.

Tabla 17.*Desviaciones estándar de talla para la edad en niñas según OMS*

Day	SD4neg	SD3neg	SD2neg	SD1neg	SD0	SD1	SD2	SD3	SD4
0	41.697	43.56	45.422	47.285	49.148	51.01	52.873	54.736	56.598
1	41.854	43.72	45.585	47.451	49.317	51.182	53.048	54.914	56.779
2	42.011	43.88	45.748	47.617	49.485	51.354	53.223	55.091	56.96
3	42.166	44.038	45.91	47.782	49.654	51.526	53.398	55.27	57.142
4	42.324	44.199	46.074	47.948	49.823	51.698	53.573	55.448	57.323
5	42.481	44.359	46.237	48.114	49.992	51.87	53.748	55.625	57.503
6	42.639	44.519	46.4	48.28	50.161	52.041	53.922	55.802	57.683
7	42.796	44.68	46.563	48.446	50.33	52.213	54.096	55.98	57.863
8	42.954	44.84	46.726	48.613	50.499	52.385	54.271	56.157	58.043
9	43.112	45.001	46.89	48.779	50.668	52.556	54.445	56.334	58.223

Nota. Obtenido de OMS (2017). *Extracto desviaciones estándar talla para la edad en niñas.* (www.who.int/tools/child-growth-standards/standards/length-height-for-age), consultado el 21 de septiembre de 2021. De dominio público.

En las tablas 18 y 19 se presentan extractos de los estándares publicados por la OMS para el seguimiento nutricional de peso para la talla en desviaciones estándar, para niños y niñas, respectivamente. El peso se presenta en kilogramos y la talla en centímetros.

Tabla 18.

Desviaciones estándar de peso para la talla en niños

Talla	SD4neg	SD3neg	SD2neg	SD1neg	SD0	SD1	SD2	SD3	SD4
45	1.71	1.877	2.043	2.23	2.441	2.68	2.951	3.261	3.571
45.1	1.722	1.89	2.057	2.245	2.458	2.698	2.971	3.283	3.595
45.2	1.734	1.903	2.072	2.261	2.474	2.716	2.991	3.305	3.618
45.3	1.747	1.916	2.086	2.276	2.491	2.734	3.011	3.326	3.642
45.4	1.759	1.929	2.1	2.292	2.508	2.752	3.031	3.348	3.666

Nota. Obtenido de OMS (2017). *Extracto desviaciones estándar peso para la talla en niños.* (<https://www.who.int/tools/child-growth-standards/standards/weight-for-age>), consultado el 21 de septiembre de 2021. De dominio público.

Tabla 19.

Desviaciones estándar de peso para la talla en niñas

Talla	SD4neg	SD3neg	SD2neg	SD1neg	SD0	SD1	SD2	SD3	SD4
45	1.737	1.902	2.066	2.252	2.461	2.698	2.967	3.275	3.584
45.1	1.749	1.915	2.081	2.267	2.478	2.716	2.988	3.298	3.609
45.2	1.761	1.928	2.095	2.283	2.495	2.735	3.008	3.321	3.633
45.3	1.773	1.941	2.109	2.298	2.512	2.753	3.029	3.343	3.658
45.4	1.785	1.954	2.123	2.314	2.529	2.772	3.049	3.366	3.683
45.5	1.797	1.967	2.138	2.329	2.546	2.791	3.07	3.389	3.708
45.6	1.809	1.98	2.152	2.345	2.563	2.809	3.09	3.412	3.733

Nota. Obtenido de OMS (2017). *Extracto desviaciones estándar peso para la talla en niñas.* (<https://www.who.int/tools/child-growth-standards/standards/weight-for-age>), consultado el 21 de septiembre de 2021. De dominio público.

Utilizando las tablas anteriores, se clasificaron las unidades de la muestra de acuerdo con el diagnóstico según las desviaciones estándar, el cual se presenta en la tabla XX.

Tabla 20.

Clasificación nutricional infantil según desviaciones estándar

Comparativa respecto a las desviaciones estándar	Talla para edad	Peso para la talla
Menor a 3 desviaciones estándar de la media	Retraso severo de crecimiento (RSC)	Emaciación severa (EMS)
Menor a 2 desviaciones estándar de la media	Retraso de crecimiento (RC)	Emaciación (EM)
Mayor a 2 desviaciones estándar de la media	Normal (sin desnutrición)	Normal (sin desnutrición)

Nota. Obtenido de OMS (2017). *Clasificación nutricional infantil según desviaciones estándar.* (<https://www.who.int/tools/child-growth-standards/standards>), consultado el 21 de septiembre de 2021. De dominio público.

Tabla 21.

Resumen de diagnóstico desviaciones estándar

Clasificación T/E	Frecuencia T/E	Clasificación P/T	Frecuencia P/T
Normal	989	Normal	1874
RC	512	EM	42
RSC	429	EMS	14

Nota. Elaboración propia, realizado con Microsoft Excel.

La tabla 21 muestra las frecuencias de los casos estudiados según la clasificación por desviaciones estándar para los indicadores de talla para la edad y peso para la talla.

3.1.2. Análisis de devianzas para la relación entre variables independientes con la presencia de desnutrición infantil

Con el objetivo de conocer la relación logística entre las variables independientes con la variable respuesta, previo a elaborar el modelo de regresión logística, se codificaron las variables explicativas en niveles dicotómicos. En la tabla 22 se presenta la codificación de variables utilizada para dicotomizar los niveles de las variables independientes, de manera que se pudiera realizar el análisis de devianza y posteriormente, el modelo de regresión logística, con un nivel de complejidad que no fuera demasiado alto, respetando el principio de parsimonia. Los niveles que se consideraron peligrosos de acuerdo con las frecuencias de desnutrición obtenidas con ambos métodos de diagnóstico se clasificaron con el valor 1 y el resto con el valor 0.

Tabla 22.

Codificación de variables por niveles binomiales

variable	Código	Valor
piso_casa	1	Tierra, madera y otro
piso_casa	0	Ladrillo, torta de cemento, granito o ladrillo cerámico
fuelle_agua	1	Chorro público, lluvia, pozo, río, camión o pipa, de otra vivienda
fuelle_agua	0	Tubería dentro de vivienda o terreno
piso_cocina	1	Tierra, madera y otro
piso_cocina	0	Ladrillo, torta de cemento, granito o ladrillo cerámico
calidad_agua	1	hervida o ninguna
calidad_agua	0	Filtrada, clorada o purificada
combustible_cocina	1	Leña

Continuación de la tabla 22.

variable	código	Valor
combustible_cocina	0	Gas o estufa eléctrica
disposicion_basura	1	servicio informal
disposicion_basura	0	servicio formal
disposicion_excretas	1	letrina o intemperie
disposicion_excretas	0	Inodoro
tipo_casa	1	palomar, rancho, vivienda improvisada, otro tipo de vivienda
tipo_casa	0	Casa Formal, Apartamento
techo	1	lámina, asbesto, teja de barro, otro
techo	0	Concreto

Nota. Elaboración propia, realizado con Microsoft Excel.

De manera similar a la codificación utilizada para dicotomizar las variables independientes, se realizó una codificación de las variables respuesta para separar los casos que presenten algún tipo de desnutrición de los casos normales, en ambos métodos de diagnóstico. La dicotomización de las variables respuesta también es necesaria por el requisito del modelo de regresión, entendiendo que es un modelo de regresión logística binomial. Los casos que presentaron un diagnóstico positivo de cualquier tipo de desnutrición al aplicar ambos métodos se codificaron con el valor 1, y el resto con el valor 0, según se muestra en las tablas 23 y 24.

Tabla 23.*Codificación dicotómica de resultados de diagnóstico Waterlow*

Método diagnóstico	Niveles	Codificación
Waterlow	DA	1
Waterlow	DCAG	1
Waterlow	DCAR	1
Waterlow	Normal	0

Nota. Elaboración propia, realizado con Microsoft Excel.

Tabla 24.*Codificación dicotómica desviaciones estándar (ZTE y ZPT)*

Método diagnóstico	Niveles	Codificación
ZTE	Normal	0
ZTE	RC	1
ZTE	RSC	1
ZPT	Normal	0
ZPT	EM	1
ZPT	EMS	1

Nota. Elaboración propia, realizado con Microsoft Excel.

A continuación, se presenta un extracto de la tabla codificada de los 1930 casos del estudio, utilizada para realizar el análisis de devianzas. Se incluyó el resultado del diagnóstico de Waterlow (W) y los 2 resultados del diagnóstico de las desviaciones estándar (ZTE y ZPT).

Tabla 25.

Extracto codificación dicotómica para análisis de devianza

Tech	Tipo casa	Piso casa	fuelle agua	Piso cocina	Calidad agua	Combustible cocina	Disposición basura	Disposiciones excretas	W	ZTE	ZPT
1	0	0	0	1	0	1	1	1	0	0	0
1	0	1	0	1	0	1	1	0	0	0	0
1	0	1	0	1	0	1	1	1	0	0	1
1	0	1	0	1	0	1	1	1	1	1	0

Nota. Elaboración propia, realizado con Microsoft Excel.

A continuación, se presenta un extracto del script ejecutado en el software Rstudio para obtener los resultados del análisis de varianza para cada una de las variables de respuesta.

```
getwd()
data_reg <- read.csv("prueba_reg_objetivo1.csv",header = TRUE,na.strings=c(""))
data_reg<- as.data.frame(unclass(data_reg), stringsAsFactors = TRUE) # Consulta de
archivo CSV y conversión de columnas a factores

modelZTE <- glm(ZTE ~.,family=binomial(link='logit'),data=data_reg # Creación de modelo
binomial de variable respuesta 'ZTE' para análisis de devianza
anova(modelZTE, test="Chisq")
>Analysis of Deviance Table
Model: binomial, link: logit
Response: ZTE
Terms added sequentially (first to last)

modelZPT <- glm(ZPT ~.,family=binomial(link='logit'),data=data_reg # Creación de modelo
binomial de variable respuesta 'ZPT' para análisis de devianza
anova(modelZPT, test="Chisq")
```

```
>Analysis of Deviance Table
Model: binomial, link: logit
Response: ZPT
Terms added sequentially (first to last)
```

```
modelW<- glm(W~.,family=binomial(link='logit'),data=data_reg # Creación de modelo
binomial de variable respuesta 'W' para análisis de devianza
anova(modelW, test="Chisq")
```

```
>Analysis of Deviance Table
Model: binomial, link: logit
Response: W
Terms added sequentially (first to last)
```

A continuación, se presentan las tablas correspondientes a los análisis de devianza para cada una de las variables respuesta al ejecutar los scripts anteriores.

Tabla 26.
Análisis de devianza variable ZTE

Variable	gl	Devianza	Gl resid.	Devianza resid.	Pr(>Chi)	Significancia
NULL			1929	2674.3		
Techo	1	10.88	1928	2663.5	0.0009734	***
piso casa	1	28.06	1927	2635.4	1.18E-07	***
piso cocina	1	0.72	1926	2634.7	0.395394	
tipo casa	1	3.04	1925	2631.7	0.081308	.
fuelle agua	1	31.7	1924	2599.9	1.80E-08	***

Continuación de la tabla 26.

calidad agua	1	2.35	1923	2597.6	0.1255807	
combustible cocina	1	0.02	1922	2597.6	0.8866217	
disposición basura	1	1.1	1921	2596.5	0.2939634	
disposiciones excretas	1	15.84	1920	2580.6	6.89E-05	***

Nota. Elaboración propia, realizado con Rstudio.

Tabla 27.

Análisis de devianza variable ZPT

Variable	gl	Devianza	Gl resid.	Devianza resid.	Pr(>Chi)	Significancia
NULL			1929	506.83		
Techo	1	0.2294	1928	506.6	0.631999	
piso casa	1	0.5184	1927	506.08	0.471517	
piso cocina	1	3.1162	1926	502.97	0.077518	.
tipo casa	1	1.385	1925	501.58	0.239251	
fuentes agua	1	7.2058	1924	494.38	0.007267	**
calidad agua	1	0.1933	1923	494.18	0.66015	

Continuación de la tabla 27.

combustible cocina	1	0.1309	1922	494.05	0.717505	
disposición basura	1	0.7649	1921	493.29	0.381797	
disposiciones excretas	1	1.2781	1920	492.01	0.258254	

Nota. Elaboración propia, realizado con Rstudio.

Tabla 28.

Análisis de devianza variable W

Variable	gl	Devianza	Gl resid.	Devianza resid.	Pr(>Chi)	Significancia
NULL			1929	2093.9		
Techo	1	10.27	1928	2083.6	0.001353	**
piso_casa	1	27.26	1927	2056.3	1.78E-07	***
piso_cocina	1	0.1	1926	2056.2	0.754731	
tipo_casa	1	1.04	1925	2055.2	0.307685	
fuente_agua	1	36.69	1924	2018.5	1.38E-09	***

Continuación de la tabla 28.

calidad_agua	1	0.15	1923	2018.3	0.697914	
combustible_cocina	1	0.05	1922	2018.3	0.819008	
disposicion_basura	1	1.72	1921	2016.6	0.189448	
disposicion_excretas	1	9.44	1920	2007.1	0.002121	**

Nota. Elaboración propia, realizado con Rstudio.

3.2. Objetivo 2. Explicación del comportamiento de las condiciones de desnutrición infantil, en relación con la información correspondiente a los hogares, para mejorar las propuestas de desarrollo social por medio de un modelo de regresión logística multivariada en el suroccidente de Guatemala

De acuerdo con los resultados obtenidos en el objetivo anterior, se presentan las variables independientes significativas para cada una de las variables respuesta, que fueron utilizadas para desarrollar los modelos de regresión logística correspondientes.

Tabla 29.*Resumen de variables significativas*

Variables respuesta	Variables independientes significativas	Significancia según análisis de devianza
ZTE	Techo	0.0009734
ZTE	tipo_casa	0.081308
ZTE	piso_casa	1.18E-07
ZTE	fuelle_agua	1.80E-08
ZTE	disposicion_excretas	6.89E-05
ZPT	fuelle_agua	0.007267
ZPT	piso_cocina	0.077518
W	techo	0.001353
W	fuelle_agua	1.38E-09
W	piso_casa	1.78E-07
W	disposición_excretas	0.002121

Nota. Elaboración propia, realizado con Rstudio.

3.2.1. Desarrollo de modelos de regresión logística para variables respuesta

A continuación, se presentan los resultados obtenidos con el desarrollo de los modelos de regresión logística para cada una de las variables respuesta.

3.2.1.1. Modelo de regresión logística para variable ZTE

A continuación, se presentan los resultados obtenidos con el desarrollo del modelo de regresión logística para la variable ZTE, utilizando el software IBM SPSS.

Tabla 30.

Bloque 0: modelo ZTE

	Observado	Pronosticado		
		ZTE		Porcentaje correcto
		.00	1.00	
Paso 0	ZTE .00	989	0	100.0
	1.00	941	0	.0
	Porcentaje global			48.8

a. En el modelo se incluye una constante.

b. El valor de corte es .400

Nota. Elaboración propia, realizado con IBM SPSS.

Tabla 31.

Variables en la ecuación modelo ZTE (bloque inicial)

	B	E.T.	Wald	gl	Sig.	Exp(B)
Paso 0 Constante	-.050	.046	1.194	1	.275	.951

Nota. Elaboración propia, realizado con IBM SPSS.

Tabla 32.

Variables que no están en la ecuación modelo ZTE

		Puntuación	Gl	Sig.
Paso 0	techo	10.781	1	.001
	tipo_casa	5.169	1	.023
	Variables piso_cocina	8.546	1	.003
	fuelle_agua	42.365	1	.000
	disposicion_excretas	45.891	1	.000
	Estadísticos globales	76.987	5	.000

Nota. Elaboración propia, realizado con IBM SPSS.

Tabla 33.

Pruebas omnibus sobre los coeficientes del modelo ZTE

		Chi cuadrado	Gl	Sig.
Paso		78.207	5	.000
Paso 1	Bloque	78.207	5	.000
	Modelo	78.207	5	.000

Nota. Elaboración propia, realizado con IBM SPSS.

Tabla 34.

Prueba de Hosmer y Lemeshow del modelo ZTE

Paso	Chi cuadrado	gl	Sig.
1	3.711	7	.812

Nota. Elaboración propia, realizado con IBM SPSS.

Tabla 35.*Tabla de contingencias para la prueba de Hosmer y Lemeshow*

	ZTE = .00		ZTE = 1.00		Total
	Observado	Esperado	Observado	Esperado	
1	105	100.109	45	49.891	150
2	148	150.081	90	87.919	238
3	131	134.992	91	87.008	222
4	34	36.046	32	29.954	66
Paso 1 5	129	136.725	138	130.275	267
6	2	2.500	3	2.500	5
7	293	280.714	282	294.286	575
8	68	66.187	99	100.813	167
9	79	81.646	161	158.354	240

Nota. Elaboración propia, realizado con IBM SPSS.

Tabla 36.*Tabla de clasificación modelo ZTE*

Tabla de clasificación^a

	Observado	Pronosticado		
		ZTE		Porcentaje correcto
		.00	1.00	
ZTE .00		384	605	38.8
Paso 1 ZTE 1.00		226	715	76.0
Porcentaje global				56.9

a. El valor de corte es .400

Nota. Elaboración propia, realizado con IBM SPSS.

Tabla 37.

Variables en la ecuación modelo ZTE

	B	E.T.	Wald	Gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
							Inferior	Superior
Techo	.176	.154	1.302	1	.254	1.192	.881	1.613
tipo_casa	.543	.296	3.351	1	.067	1.720	.962	3.075
piso_cocina	.096	.100	.917	1	.338	1.100	.905	1.338
Paso 1 ^a fuente_agua	.584	.121	23.304	1	.000	1.794	1.415	2.274
disposicion_excretas	.486	.106	21.182	1	.000	1.627	1.322	2.001
Constante	-.711	.141	25.486	1	.000	.491		

a. Variable(s) introducida(s) en el paso 1: techo, tipo casa, piso cocina, fuente agua, disposiciones excretas.

Nota. Elaboración propia, realizado con IBM SPSS.

3.2.1.2. Modelo de regresión logística para variable ZPT

A continuación, se presentan los resultados obtenidos con el desarrollo del modelo de regresión logística para la variable ZPT, utilizando el software IBM SPSS.

Tabla 38.*Bloque 0: modelo ZPT*

	Observado	Pronosticado		
		ZPT		Porcentaje correcto
		.00	1.00	
Paso 0	ZPT .00	1874	0	100.0
	ZPT 1.00	56	0	.0
	Porcentaje global			97.1

a. En el modelo se incluye una constante.

b. El valor de corte es .400

Nota. Elaboración propia, realizado con IBM SPSS.**Tabla 39.***Variables en la ecuación modelo ZPT (bloque inicial)*

	B	E.T.	Wald	Gl	Sig.	Exp(B)
Paso 0 Constante	-3.510	.136	670.090	1	.000	.030

Nota. Elaboración propia, realizado con IBM SPSS.**Tabla 40.***Variables que no están en la ecuación modelo ZPT*

	Puntuación	gl	Sig.
Paso 0 Variables fuente_agua	8.675	1	.003
piso_cocina	2.624	1	.105
Estadísticos globales	10.637	2	.005

Nota. Elaboración propia, realizado con IBM SPSS.

Tabla 41.*Pruebas ómnibus sobre los coeficientes del modelo ZPT*

	Chi cuadrado	Gl	Sig.
Paso	9.578	2	.008
Paso 1 Bloque	9.578	2	.008
Modelo	9.578	2	.008

Nota. Elaboración propia, realizado con IBM SPSS.**Tabla 42.***Prueba de Hosmer y Lemeshow del modelo ZPT*

Paso	Chi cuadrado	gl	Sig.
1	.148	2	.929

Nota. Elaboración propia, realizado con IBM SPSS.**Tabla 43.***Tabla de contingencias para la prueba de Hosmer y Lemeshow*

	ZPT = .00		ZPT = 1.00		Total
	Observado	Esperado	Observado	Esperado	
Paso 1 1	675	674.384	12	12.616	687
2	830	830.616	24	23.384	854
3	131	131.616	6	5.384	137
4	238	237.384	14	14.616	252

Nota. Elaboración propia, realizado con IBM SPSS.

Tabla 44.*Tabla de clasificación modelo ZPT*

	Observado	Pronosticado		
		ZPT		Porcentaje correcto
		.00	1.00	
Paso 1	ZPT .00	1874	0	100.0
	ZPT 1.00	56	0	.0
	Porcentaje global			97.1

a. El valor de corte es .400

Nota. Elaboración propia, realizado con IBM SPSS.**Tabla 45.***Variables en la ecuación modelo ZPT*

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
							Inferior	Superior
fuelle_agua	.783	.286	7.485	1	.006	2.187	1.248	3.831
Paso 1 ^a piso_cocina	.409	.291	1.968	1	.161	1.505	.850	2.664
Constante	-3.979	.253	247.137	1	.000	.019		

a. Variable(s) introducida(s) en el paso 1: fuente_agua, piso_cocina.

Nota. Elaboración propia, realizado con IBM SPSS.

3.2.1.3. Modelo de regresión logística para variable W

A continuación, se presentan los resultados obtenidos con el desarrollo del modelo de regresión logística para la variable W, utilizando el software IBM SPSS.

Tabla 46.

Bloque 0: modelo W

	Observado	Pronosticado		
		W		Porcentaje correcto
		.00	1.00	
Paso 0	W .00	1481	0	100.0
	1.00	449	0	.0
	Porcentaje global			76.7

a. En el modelo se incluye una constante.

b. El valor de corte es .400

Nota. Elaboración propia, realizado con IBM SPSS.

Tabla 47.

Variables en la ecuación modelo W (bloque inicial)

	B	E.T.	Wald	gl	Sig.	Exp(B)
Paso 0 Constante	-1.193	.054	490.741	1	.000	.303

Nota. Elaboración propia, realizado con IBM SPSS.

Tabla 48.

Variables que no están en la ecuación modelo W

		Puntuación	gl	Sig.
Paso 0	Techo	9.462	1	.002
	Variables			
	fuelle_agua	51.625	1	.000
	piso_casa	35.401	1	.000
	disposicion_excretas	35.993	1	.000
Estadísticos globales		86.933	4	.000

Nota. Elaboración propia, realizado con IBM SPSS.

Tabla 49.

Pruebas omnibus sobre los coeficientes del modelo W

		Chi cuadrado	Gl	Sig.
Paso		85.712	4	.000
Paso 1	Bloque	85.712	4	.000
	Modelo	85.712	4	.000

Nota. Elaboración propia, realizado con IBM SPSS.

Tabla 50.

Prueba de Hosmer y Lemeshow del modelo W

Paso	Chi cuadrado	gl	Sig.
1	1.058	6	.983

Nota. Elaboración propia, realizado con IBM SPSS.

Tabla 51.*Tabla de contingencias para la prueba de Hosmer y Lemeshow*

	W = .00		W = 1.00		Total
	Observado	Esperado	Observado	Esperado	
1	127	126.667	17	17.333	144
2	272	272.348	42	41.652	314
3	49	51.130	13	10.870	62
4	299	296.592	68	70.408	367
5	146	148.278	39	36.722	185
6	366	361.884	130	134.116	496
7	99	98.927	46	46.073	145
8	123	125.174	94	91.826	217

Nota. Elaboración propia, realizado con IBM SPSS.**Tabla 52.***Tabla de clasificación modelo W*

	Observado	Pronosticado		
		W		Porcentaje correcto
		.00	1.00	
W	.00	1358	123	91.7
Paso 1	1.00	355	94	20.90
	Porcentaje global			75.2

a. El valor de corte es .400

Nota. Elaboración propia, realizado con IBM SPSS.

Tabla 53.*Variables en la ecuación modelo W*

	B	Error estándar	Wald	gl	Sig.	Exp(B)	95% C.I. para EXP(B)		
							Inferior	Superior	
Paso 1 ^a									
techo	.111	.205	.294	1	.588	1.118	.748	1.671	
piso_casa	.445	.119	14.112	1	.000	1.561	1.237	1.970	
fuentes_agua	.683	.128	28.611	1	.000	1.979	1.541	2.542	
disposicion_excretas	.440	.135	10.585	1	.001	1.552	1.191	2.023	
Constante	-1.989	.190	109.16	1	.000	.137			
				2					
a. Variables especificadas en el paso 1: techo, piso_casa, fuentes_agua, disposicion_excretas.									

Nota. Elaboración propia, realizado con IBM SPSS.

3.3. Objetivo 3. Determinación de los factores físicos correspondientes a los hogares que son significativos en las condiciones de desnutrición infantil, y en qué niveles estos factores provocan condiciones de desnutrición, para mejorar las propuestas de desarrollo social mediante un modelo de regresión logística multivariada y el análisis de odds ratios en el suroccidente de Guatemala

Con los resultados obtenidos al desarrollar los modelos de regresión logística, se saben los factores y niveles que son estadísticamente significativos, estudiando los valores de significancia y odds ratios de las tablas de las variables de las ecuaciones de regresión.

En las tablas siguientes, se presentan los odds ratios de los niveles significativos de los factores determinantes para cada variable de respuesta, así como las probabilidades relativas de presentar las condiciones de desnutrición asociadas a cada variable de respuesta entre los niveles con efecto negativo y efecto positivo.

3.3.1. Factores y Odds Ratios de niveles significativos para la variable de respuesta ZTE

En el caso de la variable de respuesta ZTE (talla para la edad según diagnóstico de desviaciones estándar), las variables significativas fueron la fuente de agua y la disposición de excretas. También se consideró la variable de tipo de casa como significativa ya que su valor de significancia equivale a 0.067, se acerca en buena medida al valor límite.

Tabla 54.

Factores y niveles con efecto positivo variable ZTE

Niveles variables significativas	Niveles con efecto positivo según codificación	Significancia según modelo de regresión	Odds Ratio (OR)	Límite inferior OR (IC 95%)	Límite superior OR (IC 95%)
fuelle_agua	Chorro público, lluvia, pozo, río, camión o pipa, de otra vivienda	0.000	1.794	1.415	2.274
disposición_excretas	Letrina o intemperie	0.000	1.627	1.322	2.001
tipo casa	Palomar, rancho, vivienda improvisada, otro tipo de vivienda	0.067	1.720	0.962	3.075

Nota. Elaboración propia, realizado con Microsoft Excel.

3.3.2. Factores y Odds Ratios de niveles significativos para la variable de respuesta ZPT

En el caso de la variable de respuesta ZPT (peso para la talla según diagnóstico de desviaciones estándar), la variable significativa fue la fuente de agua en los niveles que no implican la disposición de tubería dentro de vivienda o terreno con una significancia de 0.006.

Tabla 55.

Factores y niveles con efecto positivo variable ZPT

Niveles variables significativas	Niveles con efecto positivo según codificación	Significancia según modelo de regresión	Odds Ratio (OR)	Límite inferior OR (IC 95%)	Límite superior OR (IC 95%)
fuentes_agua	Chorro público, lluvia, pozo, río, camión o pipa, de otra vivienda	0.006	2.187	1.248	3.831

Nota. Elaboración propia, realizado con Microsoft Excel.

3.3.3. Factores y Odds Ratios de niveles significativos para la variable de respuesta W

En el caso de la variable de respuesta W (diagnostico combinado según método de Waterlow), todas las variables utilizadas para el desarrollo del modelo de regresión resultaron significativas, a excepción del techo de la casa (fuente de agua, piso de casa y disposición de excretas).

La variable con el mayor OR fue la fuente de agua (1.979), la cual se repite también para los modelos anteriores como variable significativa.

Tabla 56.

Factores y niveles con efecto positivo variable W

Niveles variables significativas	Niveles con efecto positivo según codificación	Significancia según modelo de regresión	Odds Ratio (OR)	Límite inferior OR (IC 95%)	Límite superior OR (IC 95%)
fuentes_agua	Chorro público, lluvia, pozo, río, camión o pipa, de otra vivienda	0.000	1.979	1.541	2.542
disposición_excretas	Letrina o intemperie	0.001	1.552	1.191	2.023
piso_casa	Tierra	0.000	1.561	1.237	1.970

Nota. Elaboración propia, realizado con Microsoft Excel.

A continuación, se presentan los coeficientes de los modelos de regresión logística desarrollados con los que se construyen las ecuaciones correspondientes.

$$\ln \left[\frac{P(W)}{1-P(W)} \right] = -1.989 + 0.683X_1 + 0.445X_2 + 0.440X_3 \text{ [Ec. 23]}$$

Donde:

$P(W)$ es la probabilidad de que el individuo esté en condición de desnutrición,

X_1 corresponde al nivel del factor fuente de agua,

X_2 corresponde al nivel del factor piso de casa,

Y X_3 corresponde al nivel del factor disposición de excretas.

Ecuación para el modelo según el diagnóstico de T/E por desviaciones estándar (ZTE):

$$\ln \left[\frac{P(ZTE)}{1-P(ZTE)} \right] = -.711 + 0.543X_1 + 0.584X_2 + 0.486X_3 \text{ [Ec. 24]}$$

Donde:

$P(ZTE)$ es la probabilidad de que el individuo esté en condición de desnutrición,

X_1 corresponde al nivel del factor tipo de casa,

X_2 corresponde al nivel del factor fuente de agua,

Y X_3 corresponde al nivel del factor disposición de excretas.

Ecuación para el modelo según el diagnóstico de P/T por desviaciones estándar (ZPT):

$$\ln \left[\frac{P(ZPT)}{1 - P(ZPT)} \right] = -3.979 + 0.783X_1 \text{ [Ec. 25]}$$

Donde:

$P(ZPT)$ es la probabilidad de que el individuo esté en condición de desnutrición,

Y X_1 corresponde al nivel del factor fuente de agua.

4. DISCUSIÓN DE RESULTADOS

De acuerdo con los resultados obtenidos, se realiza la discusión de estos con análisis posteriores, explicaciones adicionales de los datos, hallazgos no esperados, limitaciones, preguntas no contestadas y futura investigación.

4.1. Análisis interno

A continuación, se presenta el análisis interno correspondiente a la discusión de variables significativas, validación y comparación de modelos desarrollados y la discusión de variables predictoras y niveles significativos.

4.1.1. Discusión de variables significativas relacionadas a infraestructura de vivienda con la presencia de desnutrición infantil

Al revisar los resultados obtenidos en la sección 3.1, resulta de importancia destacar que no todas las variables especificadas en los documentos de la OMS ejercieron un impacto directo en las variables de estudio. Sin embargo, las variables que resultaron determinantes según el análisis de devianza, se pudieron utilizar para el desarrollo de los modelos de regresión correspondientes. A continuación, se presenta una comparación general de las variables independientes en función de la cantidad de veces que fueron utilizadas para realizar los modelos de regresión y de su valor de devianza. Asimismo, se explica su efecto de acuerdo con la variable respuesta correspondiente, recordando que el indicador ZTE corresponde a diagnósticos de retraso de crecimiento, la variable ZPT corresponde a diagnósticos de peso

bajo o emaciación, y la variable W corresponde a diagnósticos combinados de acuerdo con la clasificación de cuadrantes de Waterlow.

Tabla 57.

Análisis de criticidad relativa de variables independientes

variable significativa	modelo ZTE		modelo ZPT		modelo W		criticidad
	devianza	significancia	devianza	Significancia	devianza	significancia	
piso_casa	28.060	1.18E-07	0.5184	0.471517	27.26	1.78E-07	alta
fuente_agua	31.743	1.80E-08	7.2058	0.007267	36.69	1.38E-09	muy alta
piso_cocina	0.72	0.395394	3.1162	0.077518	0.1	0.754731	media
calidad_agua	2.346	0.1255807	0.1933	0.66015	0.151	0.697914	media
disposicion_basura	1.101	0.2939634	0.7649	0.381797	1.722	0.189448	media
disposicion_excretas	15.84	6.89E-05	1.2781	0.258254	9.442	0.002121	alta
tipo_casa	3.04	0.081308	1.385	0.239251	1.04	0.307685	media
Techo	10.877	0.0009734	0.2294	0.631999	10.269	0.001353	alta

Nota. Elaboración propia, realizado con Microsoft Excel.

4.1.2. Validación y comparación de modelos de regresión logística desarrollados

A continuación, se presenta un análisis comparativo de los modelos de regresión logística desarrollados en base a los validadores aplicados y a sus valores de devianza.

4.1.2.1. Validación de modelos según devianza específica

Como primer método de validación de modelos de regresión, se realizó un análisis de devianza posterior al desarrollado en el objetivo 1, para estudiar el comportamiento de los valores de devianza relacionados a las variables incorporadas en los modelos de regresión desarrollados, de manera que se pudo conocer el porcentaje de devianza que mantuvo al reducir la cantidad de variables.

Tabla 58.

Análisis de devianza modelo ZTE

Variable	gl	Devianza	Gl resid.	Devianza resid.	Pr(>Chi)	Significancia
NULL			1929	2674.3		
Techo	1	10.877	1928	2663.5	0.0009734	***
piso_casa	1	28.06	1927	2635.4	1.18E-07	***
tipo_casa	1	3.245	1926	2632.2	0.0716322	.
fuentes_agua	1	31.743	1925	2600.4	1.76E-08	***
disposicion_excretas	1	16.527	1924	2583.9	4.80E-05	***

Nota. Elaboración propia, realizado con Rstudio.

Al analizar las devianzas obtenidas en la tabla 58, se observa que el valor total de devianza equivale a 90.452, que corresponde a una disminución del 3.6% de la devianza total obtenida en el objetivo 1, con todas las variables. Esta disminución se compensa con una reducción del 44.44% de las variables usadas en el objetivo 1. La diferencia entre las disminuciones porcentuales de devianza y cantidad de variables indica que el modelo desarrollado mantiene una buena

parte de significancia y bondad de ajuste, al mismo tiempo que reduce la complejidad de estudio.

Tabla 59.

Análisis de devianza específica modelo ZPT

Variable	gl	Devianza	Gl resid.	Devianza resid.	Pr(>Chi)	Significancia
NULL			1929	506.83		
piso_casa	1	2.7009	1928	506.55	0.100289	
fuelle_agua	1	6.8770	1927	499.29	0.008731	**

Nota. Elaboración propia, realizado con Rstudio.

Al analizar las devianzas obtenidas en la tabla 59, se observa que el valor total de devianza equivale a 9.5779, que corresponde a una disminución del 35.38% de la devianza total obtenida en el objetivo 1, con todas las variables. Esta disminución se compensa con una reducción del 77.78% de las variables usadas en el objetivo 1. La diferencia entre las disminuciones porcentuales de devianza y cantidad de variables indica que el modelo desarrollado mantiene una buena parte de significancia y bondad de ajuste, al mismo tiempo que reduce la complejidad de estudio.

Tabla 60.

Análisis de devianza específica modelo W

Variable	gl	Devianza	Gl resid.	Devianza resid.	Pr(>Chi)	Significancia
NULL			1929	2093.9		
Techo	1	10.269	1928	2083.6	0.001353	**
fuelle_agua	1	37.271	1927	2039.1	2.51E-11	***
piso_casa	1	27.260	1926	2019.1	1.03E-09	***
disposicion_excretas	1	10.912	1925	2008.1	0.0009553	***

Nota. Elaboración propia, realizado con Rstudio.

Al analizar las devianzas obtenidas en la tabla 60, se observa que el valor total de devianza equivale a 85.712, que corresponde a una disminución del 1.17% de la devianza total obtenida en el objetivo 1, con todas las variables. Esta disminución se compensa con una reducción del 55.56% de las variables usadas en el objetivo 1. La diferencia entre las disminuciones porcentuales de devianza y cantidad de variables indica que el modelo desarrollado mantiene una buena parte de significancia y bondad de ajuste, al mismo tiempo que reduce la complejidad de estudio. Este modelo es el que presenta un mejor comportamiento entre la relación porcentual de devianzas y variables.

4.1.2.2. Validación de modelos según prueba chi cuadrado

Al revisar de forma específica los resultados de la prueba ómnibus mostrada en la salida del proceso de modelo de regresión en el software IBM SPSS, que corresponde a la prueba de chi cuadrado que indica la razón de verosimilitud del modelo desarrollado frente al modelo nulo (en el que únicamente se incluye el Intercepto).

Tabla 61.*Verosimilitud según prueba de chi cuadrado modelo ZTE*

Pruebas omnibus sobre los coeficientes del modelo				
		Chi cuadrado	Gl	Sig.
Paso 1	Paso	78.207	4	.000
	Bloque	78.207	4	.000
	Modelo	78.207	4	.000

Nota. Elaboración propia, realizado con IBM SPSS.

Al estudiar los datos de la tabla obtenida correspondiente a las pruebas ómnibus del modelo de regresión logística desarrollado para la variable ZTE, se observa un valor de chi cuadrado equivalente a 78.207 con una significancia menor a 0.001, lo que indica que el modelo presenta una bondad de ajuste estadísticamente mejor a la del modelo nulo.

Tabla 62.*Verosimilitud según prueba de chi cuadrado modelo ZPT*

Pruebas ómnibus sobre los coeficientes del modelo				
		Chi cuadrado	Gl	Sig.
Paso 1	Paso	9.578	2	.008
	Bloque	9.578	2	.008
	Modelo	9.578	2	.008

Nota. Elaboración propia, realizado con IBM SPSS.

Al estudiar los datos de la tabla obtenida correspondiente a las pruebas ómnibus del modelo de regresión logística desarrollado para la variable ZPT, se

observa un valor de chi cuadrado equivalente a 9.578 con una significancia menor a 0.01, lo que indica que el modelo presenta una bondad de ajuste estadísticamente mejor a la del modelo nulo.

Tabla 63.

Verosimilitud según prueba de chi cuadrado modelo W

Pruebas ómnibus sobre los coeficientes del modelo				
		Chi cuadrado	Gl	Sig.
Paso 1	Paso	85.712	5	.000
	Bloque	85.712	5	.000
	Modelo	85.712	5	.000

Nota. Elaboración propia, realizado con IBM SPSS.

Al estudiar los datos de la tabla obtenida correspondiente a las pruebas ómnibus del modelo de regresión logística desarrollado para la variable W, se observa un valor de chi cuadrado equivalente a 85.712 con una significancia menor a 0.001, lo que indica que el modelo presenta una bondad de ajuste estadísticamente mejor a la del modelo nulo.

4.1.2.3. Validación de modelos según estadístico de Wald

En la siguiente tabla se muestra una comparativa del estadístico de Wald promedio por modelo obtenido en los resultados de los modelos desarrollados de regresión logística en el software IBM SPSS.

Tabla 64.

Análisis del estadístico de Wald

Modelo	Promedio estadístico Wald
ZTE	10.011
ZPT	4.727
W	13.4

Nota. Elaboración propia, realizado con IBM SPSS.

Al estudiar la tabla 64 se observa que todos los modelos desarrollados presentan un promedio para el estadístico de Wald significativamente distinto del valor 1, lo que indica que los modelos tienen una significancia general adecuada. Sin embargo, la variable piso cocina del modelo ZTE se encuentra cerca del valor 1 por lo que se descartarán en los siguientes análisis.

4.1.2.4. Validación de modelos según criterio de información de Akaike (AIC)

Para validar que los modelos desarrollados presentan un equilibrio entre el ajuste de los datos y una complejidad baja se utilizó el criterio de información de Akaike. Los valores del AIC se calcularon con todas las variables y posteriormente con las variables incluidas en los modelos para conocer el porcentaje de decremento. Asimismo, se compararon los criterios entre los modelos desarrollados.

Tabla 65.*Validación de modelos de regresión según AIC*

modelo	AIC variables completas	AIC variables incluidas en RL	% reducción
ZTE	2600.6	2595.9	0.18%
ZPT	512.01	505.3	1.31%
W	2027.1	2018.1	0.44%

Nota. Elaboración propia, realizado con IBM SPSS.

Según se observa en la tabla LXVII, el modelo con el mejor AIC es el ZPT, lo que es comprensible al recordar que consiste en 2 variables regresoras. Asimismo, fue el que mayor porcentaje de reducción presentó en comparación con el modelo completo. El segundo mejor modelo tomando en cuenta este criterio fue el modelo W, seguido por el modelo ZTE.

4.1.2.5. Validación de modelos según prueba de Hosmer Lemeshow

Como resultado del cálculo de regresión logística en el software SPSS, se analizaron los valores obtenidos en la prueba de Hosmer Lemeshow.

Tabla 66.*Validación de modelos prueba de Hosmer Lemeshow*

Modelo	Chi cuadrado	Gl	Sig.
ZTE	3.711	7	.812
ZPT	.148	2	.929
W	1.058	6	.983

Nota. Elaboración propia, realizado con IBM SPSS.

Al analizar los resultados de las pruebas de Hosmer Lemeshow aplicada a los modelos de regresión desarrollados, se observa que los 3 modelos presentan un p-valor cercano a 1, lo que indica que no existe evidencia suficiente para concluir que la bondad de ajuste de los modelos es insuficiente. Los modelos ZPT y W presentan valores mayores a 0.9 y el modelo ZTE presenta un valor mayor a 0.8.

4.1.3. Discusión de variables y niveles significativos como causas de condiciones de desnutrición infantil

Al desarrollar y analizar los resultados de los modelos de regresión para las 3 variables de respuesta con las pruebas de validación y bondad de ajuste entre modelos, variables y niveles, es posible generar un resumen que indica los factores que resultan significativos para cada variable de respuesta y su probabilidad asociada de presentar una condición de desnutrición infantil. Asimismo, se hace posible realizar análisis específicos en los factores significativos para conocer el comportamiento de la variable respuesta al mover el valor entre los niveles con efecto negativo y los niveles con efecto positivo.

4.1.3.1. Comparación de variables significativas según probabilidad de presencia de desnutrición

En la siguiente tabla se muestra una comparativa del cambio en probabilidades de presencia de desnutrición a mover los niveles de las variables de entre efectos negativos y positivos.

Tabla 67.

Cambio en probabilidades con niveles de efecto positivo

Modelo	variable significativa	aumento de probabilidad de presencia de desnutrición entre niveles
ZPT	fuelle_agua	218%
W	fuelle_agua	97.90%
ZTE	fuelle_agua	79.40%
ZTE	disposición_excretas	62.70%
W	piso_casa	56.10%
W	disposición_excretas	55.20%

Nota. Elaboración propia, realizado con IBM SPSS.

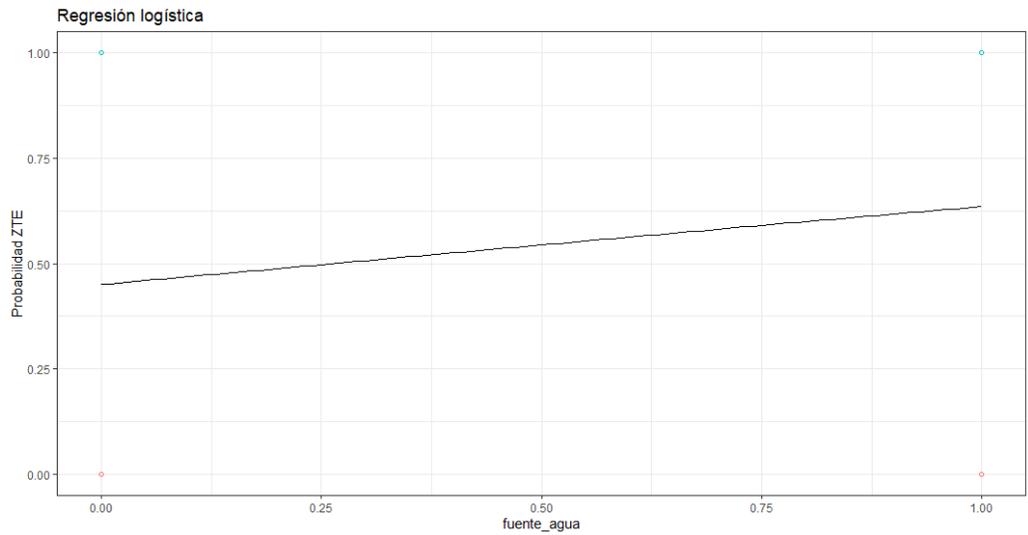
Al analizar las probabilidades presentadas en la tabla 67, se observa que la variable fuente de agua tiene el mayor efecto en todas las formas de desnutrición de todos los diagnósticos, por lo que debería ser el enfoque principal en el desarrollo de propuestas de mejora, al analizar la problemática de desnutrición infantil en el área de estudio. Sin embargo, las variables de disposición de excretas y piso de casa también presentan un nivel suficiente de impacto, como para considerarlas en tales propuestas.

4.1.3.2. Análisis de comportamiento específico de variables de respuesta por factor significativo

En las siguientes figuras se muestra cómo se incrementa la probabilidad de desnutrición entre el cambio de niveles de cada variable significativa en específico.

Figura 5.

Aumento de probabilidad de desnutrición ZTE según variable fuente_agua

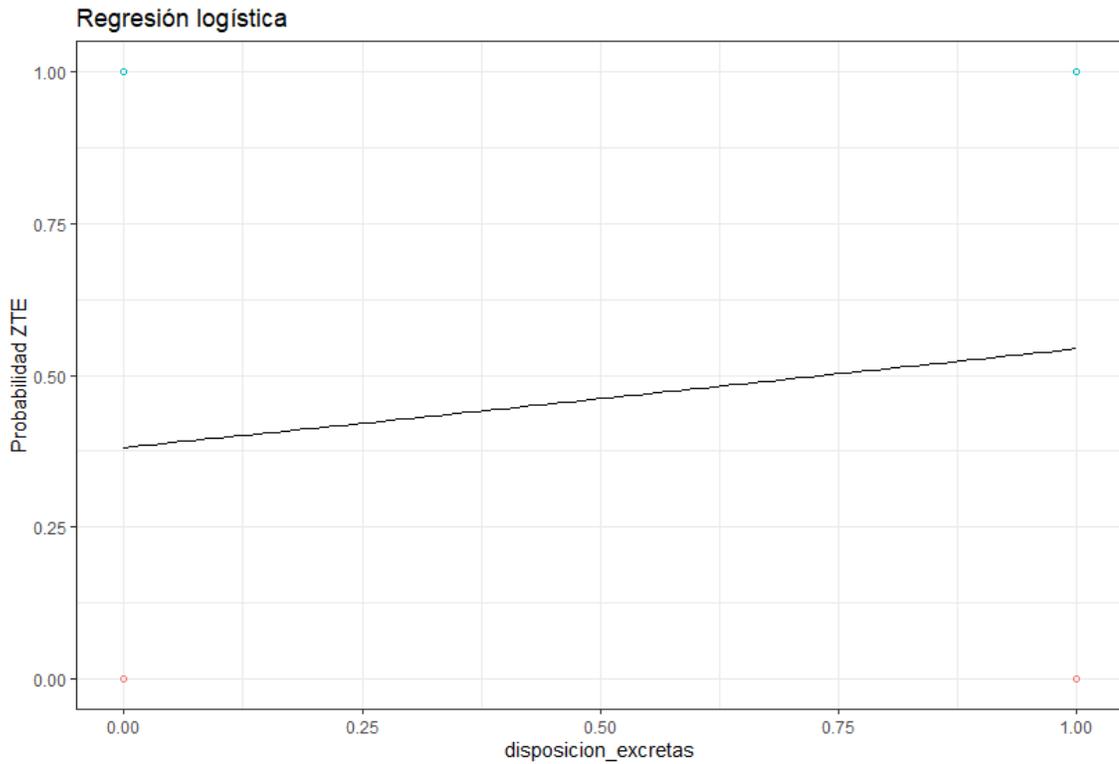


Nota. Elaboración propia, realizado con Rstudio.

Al analizar la figura 5 se observa que la probabilidad de desnutrición según la variable ZTE en el factor de la fuente de agua aumenta hasta llegar cerca del 60 % en la muestra estudiada.

Figura 6.

Aumento de probabilidad de desnutrición ZTE según variable disposición_excretas

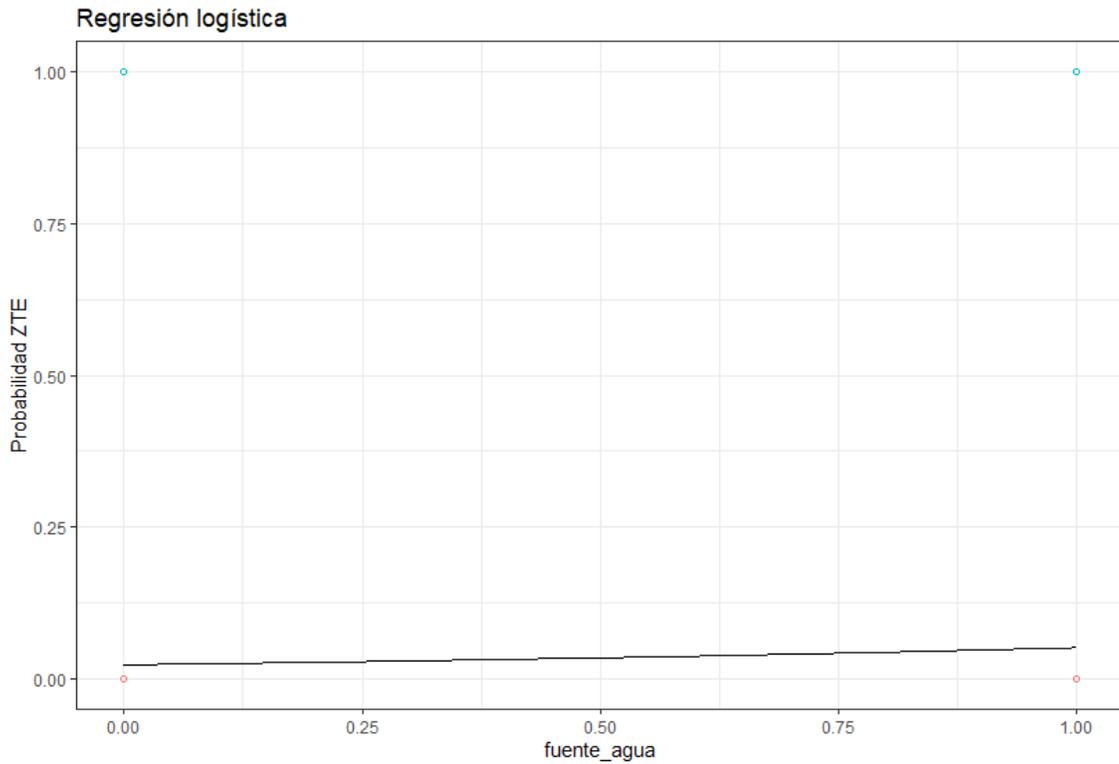


Nota. Elaboración propia, realizado con Rstudio.

Al analizar la figura 6 se observa que la probabilidad de desnutrición según la variable ZTE en el factor de la disposición de excretas aumenta hasta llegar arriba del valor de 50 %.

Figura 7.

Aumento de probabilidad de desnutrición ZPT según variable fuente_agua

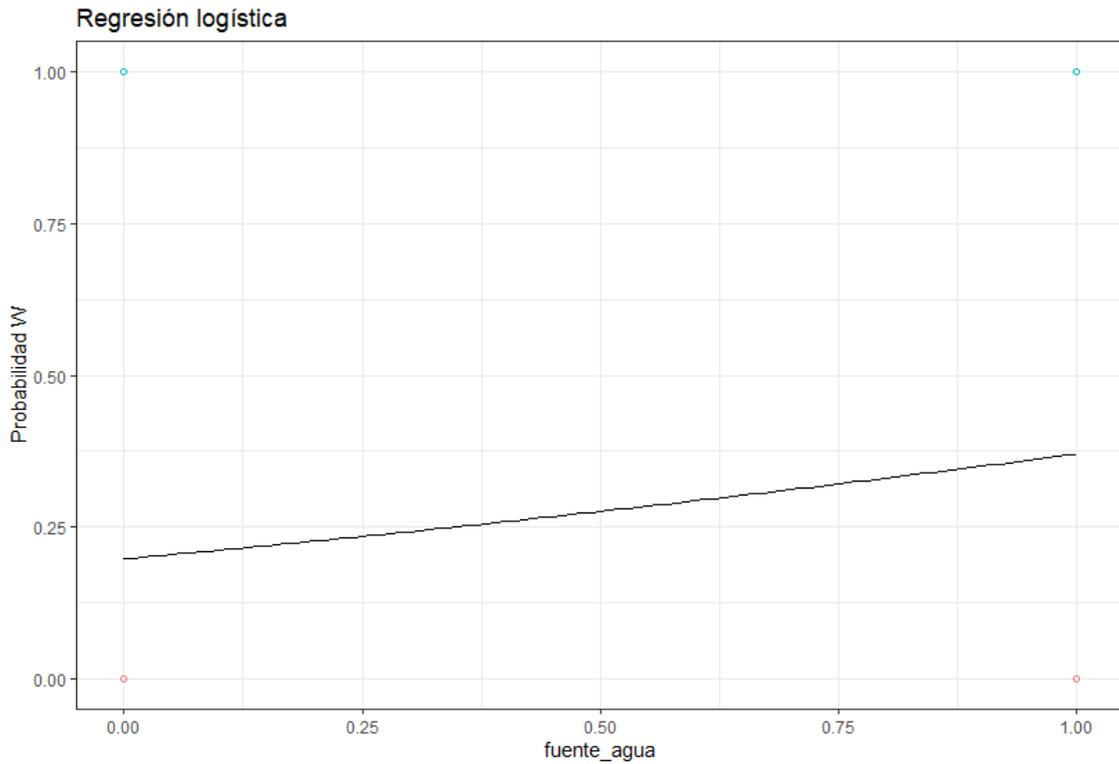


Nota. Elaboración propia, realizado con Rstudio.

Al analizar la figura 7 se observa que la probabilidad de desnutrición según la variable ZPT en el factor de la fuente de agua no aumenta hasta el valor de corte, lo que se debe a la baja cantidad de casos clasificados en este diagnóstico. Sin embargo, el aumento observado equivale aproximadamente a un 15 %.

Figura 8.

Aumento de probabilidad de desnutrición W según variable fuente_agua

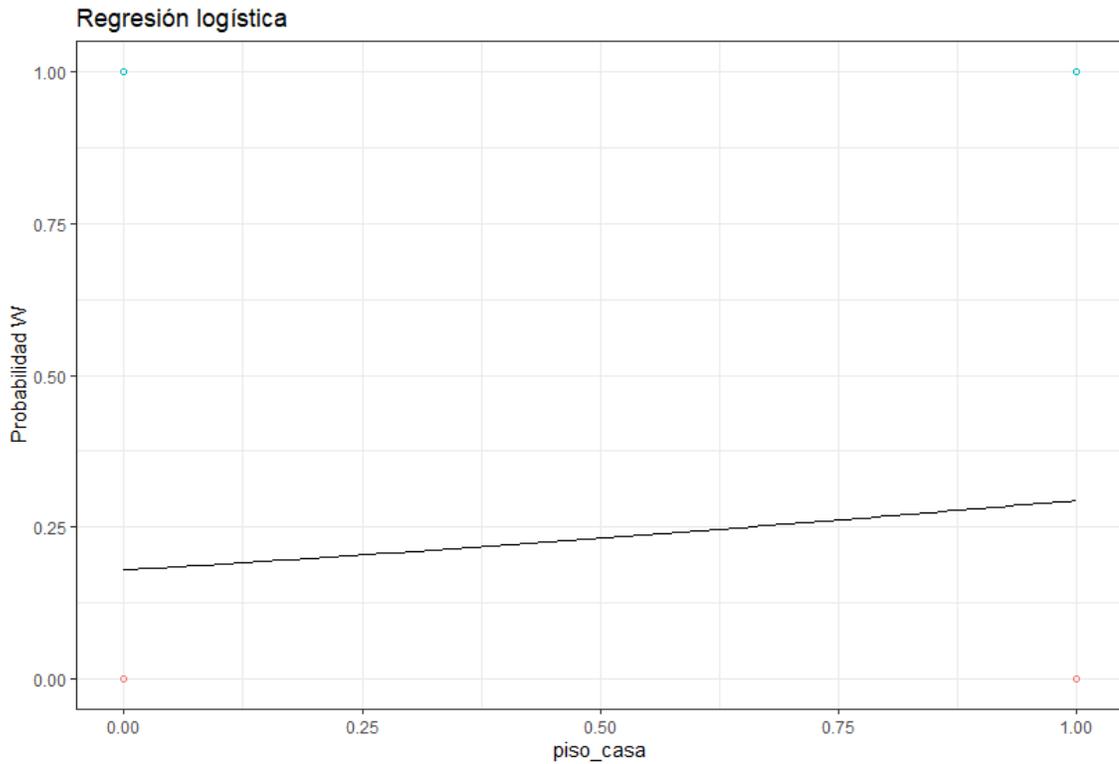


Nota. Elaboración propia, realizado con IBM SPSS.

Al analizar la figura 8 se observa que la probabilidad de desnutrición según la variable W en el factor de la fuente de agua crece aproximadamente un 20 % en la muestra estudiada, hasta llegar aproximadamente al 40 %.

Figura 9.

Aumento de probabilidad de desnutrición W según variable piso_casa

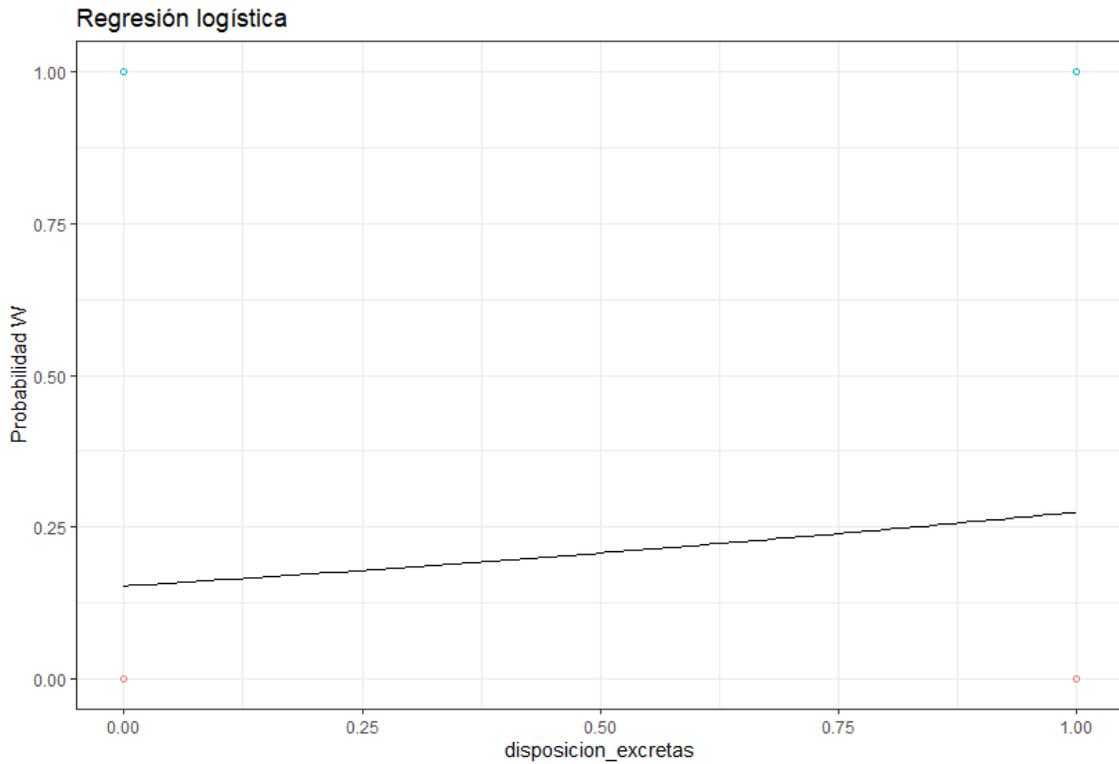


Nota. Elaboración propia, realizado con Rstudio.

Al analizar la figura 9 se observa que la probabilidad de desnutrición según la variable W en el factor del material de piso crece aproximadamente un 20 % en la muestra estudiada, hasta llegar sobre el valor del 25 %.

Figura 10.

Aumento de probabilidad de desnutrición W según variable disposición_excretas



Nota. Elaboración propia, realizado con Rstudio.

Al analizar la figura 10 se observa que la probabilidad de desnutrición según la variable W en el factor de la disposición de excretas crece aproximadamente un 15 % en la muestra estudiada, hasta llegar sobre el valor del 25 %.

4.2. Análisis externo

Al comparar los resultados obtenidos en el presente estudio con los estudios revisados en el marco referencial, se pueden analizar las coincidencias y diferencias encontradas, las cuales se resumen a continuación.

En comparación con el estudio realizado por Álvarez (2019), se llegó a la misma conclusión de relación entre la desnutrición por antropometría y el inadecuado saneamiento básico de la vivienda (desagüe, eliminación de residuos sólidos, y agua). Sin embargo, la cantidad de muestra estudiada es menor a la utilizada en el presente estudio, y la cantidad de variables analizadas es mucho mayor, tomando en cuenta que el presente estudio no considera accesibilidad económica, bajo peso al nacer, escolaridad de los padres y desnutrición materna, por motivos de accesibilidad y complejidad de factores. Debido a esto, queda como una directriz a futuro la inclusión de estas variables en una muestra similar a la del estudio realizado.

Al analizar el estudio realizado por López (2015), cuyos métodos de análisis fueron el análisis de varianza, análisis de componentes principales, cadenas de Márkov y el modelo multivariado de estimación bayesiana, se identificaron como variables explicativas los servicios públicos y el tipo de vivienda, coincidiendo con el presente estudio. Sin embargo, también se identificaron otros factores como variables significativas, incluyendo el género. La inclusión de género como variable explicativa de condiciones de desnutrición contrasta con el presente estudio debido a que los métodos de diagnóstico aceptados se basan en medidas de referencia separadas por género. No obstante, resulta de interés el análisis específico de este factor.

En el marco de los estudios realizados en Guatemala, al comparar el presente estudio con el realizado por López Bonilla (2016), se coincide en la falta de higiene de vivienda como causa de estados de desnutrición, a nivel de variable generalizada. Sin embargo, el porcentaje hallado en ese estudio de niños con desnutrición según el indicador del peso para la talla de acuerdo con el método de desviaciones de la OMS es significativamente mayor al encontrado en el presente estudio, lo que podría explicarse a la condición hospitalaria de los niños analizados. En adición a esto, el estudio no considera el indicador de la talla para la edad, que de acuerdo con el presente estudio y a los datos oficiales de desnutrición reportados, constituye el mayor porcentaje de niños en condición de desnutrición y la mayor oportunidad para desarrollar estrategias de mejora.

Al revisar el estudio realizado por Alvarado et al. (2013), continuando en el marco de estudios nacionales, en el que se describieron las condiciones sociales de niños con desnutrición respecto a la vivienda, el 69.4 % contaban con agua intradomiciliaria, mientras que el 42.1 % tuvieron letrina, 33.3 % sanitario, y 52 % piso de tierra. Estos resultados se asemejan a los porcentajes de la muestra analizada en el presente estudio. Sin embargo, al ser un acercamiento únicamente de descripción porcentual, no se presenta evidencia de la relación estadística entre los factores descritos y las condiciones de desnutrición.

Para finalizar el análisis de contraste de estudios nacionales, se revisó el estudio realizado por la ASB (2016). Los datos obtenidos se asemejan en gran medida a los obtenidos al describir la muestra del presente estudio, en cuanto a prevalencia de desnutrición crónica. Asimismo, destacan los altos porcentajes descriptivos de casos de desnutrición en los niveles de factores explicativos identificados como peligrosos en el presente estudio. Sin embargo, al igual que el estudio realizado por Alvarado et al. (2013), constituye información descriptiva

por lo que no se tiene certeza estadística de la relación de estos factores con las condiciones de desnutrición identificadas.

CONCLUSIONES

1. Se determinó la medida en que se relacionan estadísticamente las condiciones de desnutrición infantil con las características físicas y energéticas de los hogares, utilizando análisis de devianza. Con el análisis se lograron identificar 5 variables con efecto significativo en alguna de las 3 variables de respuesta, de acuerdo con los métodos de diagnóstico implementados.
2. Se lograron desarrollar modelos de regresión logística binomial para cada una de las variables de respuesta incluidas en el estudio, aplicando pruebas de validación de bondad de ajuste y complejidad. Los 3 modelos demostraron ser útiles para explicar el comportamiento de las variables respuesta con valores adecuados.
3. Se determinaron los factores físicos determinantes para explicar y pronosticar condiciones de desnutrición infantil utilizando el análisis de devianza y los odds ratios obtenidos con los modelos de regresión desarrollados. Los factores y niveles determinantes identificados fueron la fuente de agua cuando no es tubería dentro de vivienda, el piso de tierra de la casa y la disposición de excretas que no esté constituida por inodoro instalado.
4. Se desarrollaron modelos multivariados de regresión logística, que pueden ser usados para mejorar las propuestas de desarrollo social en el suroccidente de Guatemala, relacionadas a la desnutrición infantil.

RECOMENDACIONES

1. Se sugiere utilizar la información obtenida con los modelos desarrollados en el presente estudio, para proponer y ejecutar estrategias de desarrollo social en el suroccidente de Guatemala, relacionadas a la desnutrición infantil.
2. Se sugiere ampliar los análisis de devianza utilizados para identificar las medidas de relación estadística entre los factores explicativos estudiados y los indicadores de desnutrición, utilizando factores adicionales descritos en estudios previos.
3. Se sugiere utilizar a futuro los modelos de regresión logística desarrollados para predecir y explicar las condiciones de desnutrición infantil, en función de los factores explicativos considerados, con muestras y condiciones diversas.
4. Se hace necesario dar seguimiento a los factores explicativos determinados como significativos, proponiendo análisis específicos, con codificación modificada, segmentación de niveles adicionales, y otros cambios que se consideren necesarios.

REFERENCIAS

- Alvarado, M., Hernández, G., Beltrán, E., Contreras, J. (2013). *CONTEXTO SOCIAL, FAMILIAR Y DE SALUD EN MENORES DE 5 AÑOS CON DESNUTRICIÓN*. (Tesis de licenciatura). Universidad de San Carlos de Guatemala, Guatemala.
- Álvarez, L. (2019). Desnutrición infantil, una mirada desde diversos factores. *Revista Investigación Valdizana* 13(1), 15-26.
- Arias, F. (2012). El proyecto de investigación. Introducción a la metodología científica. (5ª ed.). Caracas, Venezuela: Episteme.
- ASB (Arbeiter-Samariter-Bund). (2016). Evaluación del estado nutricional y de la seguridad alimentaria de poblaciones vulnerables del corredor seco centroamericano. Guatemala: Autor
- Baños, R., Torrado Fonseca, M., Reguant, M. (julio, 2019). Análisis de regresión lineal múltiple con SPSS: un ejemplo práctico. *Revista d'Innovació i Recerca en Educació* (vol. 12 no. 2), 1-10.
- Barón, J. y Téllez, F. (diciembre, 2004) *Apuntes de Bioestadística* presentado en el tercer ciclo de la carrera en Ciencias de la Salud y Medicina, Universidad de Málaga, España.

- Calderón, D., Navarrete, C., Diaz, J. (noviembre, 2015). Ajustes de distribuciones probabilísticas para la variable temperatura media multianual para el departamento de Boyacá (Colombia). *Revista Ingeniería y Región* (vol. 14 no. 2, 125-142).
- Castillo, M. (2017). *EVALUACIÓN DEL COSTO BENEFICIO DEL CENTRO DE RECUPERACIÓN NUTRICIONAL DEL ÁREA IXIL, DEPARTAMENTO DE QUICHÉ*. (Tesis de maestría). Universidad de San Carlos de Guatemala, Guatemala.
- Cerda, J., Vera, C., Rada, G. (octubre, 2013). Odds ratio: aspectos teóricos y prácticos. *Revista médica de Chile* (vol. 141 no. 10), 1329-1335.
- Díez, A. Marrodán, M. (2018) *La desnutrición infantil en el mundo: herramientas para su diagnóstico, (1ra Ed.)* Edición Punto Didot Educación.
- Dobson, A. J., Barnett, A. G. 2008. *An Introduction to Generalized Linear Models*. 3rd ed. Chapman; Hall/CRC Press.
- Fiuza, D., Rodríguez, J. (diciembre, 2000). La regresión logística: una herramienta versátil. *Revista nefrología* (vol. 20 no. 6, 477-565).
- García Pérez, R., García Pino, G., González Ballester, D., García Moreno, R. (agosto, 2010). Modelo de regresión logística para estimar la dependencia según la escala de Lawton y Brody. *Revista Medicina de Familia – SEMERGEN* (vol. 36 no. 7), 365-371.

González, P., Vega, L. (1994). Condiciones sociodemográficas y estado nutricional de niños menores de un año en áreas periféricas de Guadalajara, México. *Revista saúde pública* 28(4), 268-276.

Hispano Portuguesa SPSS (s.f.) SPSS Guía para el análisis de datos, en análisis de regresión lineal: el procedimiento regresión lineal (Cap. 18 pp 1-69) Madrid, España.

Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. *Statistics in Medicine* 1997; Vol 16: 965-980.

ICEFI (2019). *Desnutrición crónica infantil en Guatemala: una tragedia que el debate político no debe evadir*. Guatemala: Autor.

INCAP (2012). *Guía técnica para la estandarización en procesamiento, análisis e interpretación de indicadores antropométricos según los Patrones de Crecimiento de la OMS para menores de 5 años*. Guatemala: Autor.

Islas, F. y Vela, F. (junio, 2010) *Regresión lineal multivariada, presentada en los Talleres de verano en métodos de investigación en población y territorio*. Universidad Autónoma Metropolitana, Unidad Xochimilco, México D.F.

Kragel, E., Merz, A., Flood, D., Haven, K. (2020). Risk Factors for Stunting in Children under the Age of 5 in Rural Guatemalan Highlands. *Annals of Global Health*, 86(1): 8, 1-5.

Kurtner, M. Nachtsheim, C., Neter, J. & Li, W. (2005) *Applied Linear Statistical Models, (5ta Ed.)* New York: McGraw-Hill Irwin.

- Lera, L., Albala, C., Santos, J., Arroyo, P., García, G., & Bunout, Daniel. (2007). Estudio multivariado de la evolución de las dimensiones corporales en chilenas mayores de 60 años. *Archivos Latinoamericanos de Nutrición*, 57(3), 219-223.
- Levine y Rubin (2004). *Estadística para Administración y Economía (7ma edición)*. México: Pearson Education.
- Lind, D., Marchal, W. y Wathen, S. (2008) *Estadística Aplicada a los Negocios y Economía* (13 ed.) México: McGraw-Hill Interamericana.
- López Bonilla, D. (2016). *CLASIFICACIÓN DEL ESTADO NUTRICIONAL MEDIANTE LA VALORACIÓN ANTROPOMÉTRICA EN NIÑOS/NIÑAS HOSPITALIZADOS*. (Tesis de maestría). Universidad de San Carlos de Guatemala, Guatemala.
- López, Y. (2015). *ANTROPOMETRÍA CRANEOFACIAL EN NIÑOS DE 0 A 4 AÑOS- UNA PERSPECTIVA BAYESIANA*. (Tesis de maestría). Universidad Nacional de Colombia, Colombia.
- Márquez-González, H., García-Sámano, V., Caltenco-Serrano, M., García Villegas, E., Márquez-Flores, H., Villa-Romero, A. (2012). Clasificación y evaluación de la desnutrición en el paciente pediátrico. *El Residente*; Volumen VII Número 2: 59-69

Mejía, S., Linares, I., Aguilar, J. (septiembre, 2006) *DIAGNÓSTICO DE LA PROBLEMÁTICA DE VIVIENDA EN LA REGIÓN V DE GUATEMALA*, Universidad de San Carlos de Guatemala, Guatemala.

Morales, A. (2012) *Estadística y probabilidades*, presentado en la Universidad Católica de la Santísima Concepción, Chile.

Noste, M (2013) *Apuntes de regresión lineal* presentado en la Facultad de Ciencias Exactas y Naturales para la carrera de Especialización en estadística para Ciencias de la Salud. Buenos Aires, Argentina.

Ochoa-Díaz-López H, García-Parra E, Flores-Guillén E, García-Miranda R, Solís-Hernández R. (2017). Evaluación del estado nutricional en menores de cinco años: concordancia entre índices antropométricos en población indígena de Chiapas (México). *Nutrición Hospitalaria*; Volumen 34:820-826

OMS (1990). *Principios de higiene de la vivienda*. España: Autor.

Palomo, J. (julio, 2011). *Regresión lineal múltiple*. Trabajo presentado en la Escuela Universitaria de Arquitectura Técnica de la Universidad Politécnica de Madrid.

Rioja, L., Almazán, A., Mañas, B. 2013. *Regresión Logística: Fundamentos y aplicación a la investigación sociológica*. Distribuido bajo Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0

Santiago, J., Magalhães, T., Sampaio, R. (mar, 2015). Asociación entre exceso de peso y características de adultos jóvenes escolares: contribución para el cuidado de enfermería. *Revista Latino-Am Enfermagem* 23(2), 250-258.

Spiegel, M. y Stephens, L. (2009) *Estadística* (4ta ed.) México: McGraw-Hill

Tovar, S. Navarro, J. Fernández, M. (junio, 1997). Evaluación del Estado Nutricional en Niños Conceptos Actuales. *Honduras Pediátrica*, volumen XVIII, No. 2.

Valdivia, S., Bruno, A., Romero, M. (abril, 2017). Características antropométricas maternas y del lactante, correlacionado a la concentración proteica del calostro y leche madura: un estudio longitudinal. *Revista chilena de obstetricia y ginecología* (vol. 82 no. 2), 139-146.

Valle, D. (2018). *Modelo de regresión lineal multivariado aplicado a la relación entre variables químicas del Clinker y variables físicas del cemento*. (Tesis de maestría). Universidad de San Carlos de Guatemala, Guatemala.

APÉNDICES

Apéndice 1.

Tabla de cantidad de niños del estudio por departamento y municipio

municipio	Departamento	niños del estudio
Chajul	Quiché	253
Chiantla	Huehuetenango	213
San Juan Cotzal	Quiché	209
Todos Santos Cuchumatán	Huehuetenango	163
Sacapulas	Quiché	142
Cunén	Quiché	139
Jacaltenango	Huehuetenango	137
Uspantán	Quiché	121
San Lorenzo	San Marcos	77
San Miguel Ixtahuacán	San Marcos	63
La Libertad	Huehuetenango	53
San Sebastián Huehuetenango	Huehuetenango	46
Unión Cantinil	Huehuetenango	45
Nebaj	Quiché	38
San Antonio Huista	Huehuetenango	33
La Democracia	Huehuetenango	30
San Pablo	San Marcos	25
Tajumulco	San Marcos	24
El Rodeo	San Marcos	21
San Rafael Pie de la Cuesta	San Marcos	21
Nuevo Progreso	San Marcos	21
Petatán	Huehuetenango	15
Concepción Huista	Huehuetenango	12
San Marcos	San Marcos	3
Canillá	Quiché	2
Huehuetenango	Huehuetenango	2
Sibinal	San Marcos	1

Nota. Elaboración propia, realizado con Microsoft Excel.

Apéndice 2.

Matriz de consistencia

“Modelo de regresión multivariado aplicado a los indicadores nutricionales infantiles, para mejorar las propuestas de desarrollo social en el suroccidente de Guatemala”

PROBLEMA	Preguntas de investigación	OBJETIVOS	METODOLOGÍA	RESULTADOS	CONCLUSIONES	RECOMENDACIONES
<p>Problema General No existe un modelo de regresión aplicado a los indicadores nutricionales infantiles, que permita mejorar las propuestas de desarrollo social en el suroccidente de Guatemala.</p>	<p>Pregunta principal de investigación: ¿Qué modelo de regresión permite explicar el comportamiento de los indicadores nutricionales infantiles, mejorar las propuestas de desarrollo social en el suroccidente de Guatemala?</p>	<p>Objetivo General: Desarrollar un modelo de regresión multivariado aplicado a los indicadores nutricionales infantiles, para mejorar las propuestas de desarrollo social en el suroccidente de Guatemala.</p>	<p>Tipo de Investigación: Descriptiva, transversal, correlacional</p> <p>Diseño de investigación: No experimental, retrospectivo, descriptivo, correlacional</p>	<p>Resultados generales: Se calculó la relación de devianzas encontrando 5 variables significativas, con los cuales se desarrollaron 3 modelos de regresión logística, de los cuales se obtuvieron 3 variables con niveles asociados a desnutrición.</p>	<p>Conclusión General: Se desarrollaron modelos multivariados de regresión logística, que pueden ser usados para mejorar las propuestas de desarrollo social en el suroccidente de Guatemala, relacionadas a la desnutrición infantil.</p>	<p>Recomendación General: Se sugiere utilizar la información obtenida con los modelos desarrollados en el presente estudio, para proponer y ejecutar estrategias de desarrollo social en el suroccidente de Guatemala, relacionadas a la desnutrición infantil.</p>
<p>Problemas Específicos: No se ha determinado cuál es la medida en la que se relacionan estadísticamente las condiciones de desnutrición infantil con las características físicas y energéticas de los hogares, para mejorar las propuestas de desarrollo social en el suroccidente de Guatemala.</p>	<p>Preguntas complementarias de investigación</p> <p>a) ¿Cuál es la medida en la que se relacionan estadísticamente las condiciones de desnutrición infantil con las características físicas y energéticas de los hogares, para mejorar las propuestas de desarrollo social en el suroccidente de Guatemala?</p>	<p>Objetivos Específicos: Determinar cuál es la medida en la que se relacionan estadísticamente las condiciones de desnutrición infantil con las características físicas y energéticas de los hogares, para mejorar las propuestas de desarrollo social en el suroccidente de Guatemala, utilizando análisis de devianza.</p>	<ul style="list-style-type: none"> • Población: niños de 0 a 5 años del área rural del suroccidente de Guatemala • Muestra Proporcional • Técnica Análisis multivariado de datos 	<p>Resultados específicos: Se obtuvieron 5 variables predictoras con una relación de devianza significativa, correspondientes a la infraestructura de las viviendas estudiadas. Las variables son el techo, el tipo de casa, el piso de casa, la fuente de agua y la disposición de excretas con un valor p menor a 0.05.</p>	<p>Conclusiones específicas: Se determinó la medida en que se relacionan estadísticamente las condiciones de desnutrición infantil con las características físicas y energéticas de los hogares, utilizando análisis de devianza. Con el análisis se lograron identificar 8 variables con efecto significativo en alguna de las 3 variables de respuesta, de acuerdo con los métodos de diagnóstico implementados.</p>	<p>Recomendaciones específicas: Se sugiere ampliar los análisis de devianza utilizados para identificar las medidas de relación estadística ente los factores explicativos estudiados y los indicadores de desnutrición, utilizando factores adicionales descritos en estudios previos.</p>

Continuación del apéndice 2.

No se ha explicado el comportamiento de los indicadores nutricionales infantiles, utilizando la información relacionada a los hogares (en caso existiera correlación entre ambos grupos de variables), para mejorar las propuestas de desarrollo social en el suroccidente de Guatemala.

b) ¿Cómo se puede explicar el comportamiento de los indicadores nutricionales infantiles, utilizando la información relacionada a los hogares (en caso existiera correlación entre ambos grupos de variables), para mejorar las propuestas de desarrollo social en el suroccidente de Guatemala?

Explicar el comportamiento de los indicadores nutricionales infantiles, utilizando la información relacionada a los hogares (en caso existiera correlación entre ambos grupos de variables), para mejorar las propuestas de desarrollo social en el suroccidente de Guatemala, por medio de un modelo de regresión logística multivariada.

Se desarrollaron 3 modelos de regresión logística binomial obteniendo las siguientes ecuaciones:

$$\ln \left(\frac{P_{12}}{1-P_{12}} \right) = -0.711 + 0.543X_1 + 0.594X_2 + 0.486X_3$$

$$\ln \left(\frac{P_{22}}{1-P_{22}} \right) = -3.979 + 0.783X_1$$

$$\ln \left(\frac{P_{32}}{1-P_{32}} \right) = -1.009 + 0.683X_1 + 0.445X_2 + 0.448X_3$$

Se lograron desarrollar modelos de regresión logística binomial para cada una de las variables de respuesta incluidas en el estudio, aplicando pruebas de validación de bondad de ajuste y complejidad. Los 3 modelos demostraron ser útiles para explicar el comportamiento de las variables respuesta con valores adecuados.

Se sugiere utilizar a futuro los modelos de regresión logística desarrollados para predecir y explicar las condiciones de desnutrición infantil, en función de los factores explicativos considerados, con muestras y condiciones diversas.

No se han determinado cuáles son los factores físicos correspondientes a los hogares que son significativos en los indicadores nutricionales infantiles, y en qué niveles estos factores provocan valores nutricionales bajos o deficientes, para mejorar las propuestas de desarrollo social en el suroccidente de Guatemala.

c) ¿Cuáles son los factores físicos correspondientes a los hogares que son significativos en los indicadores nutricionales infantiles, y en qué niveles estos factores provocan valores nutricionales bajos o deficientes, para mejorar las propuestas de desarrollo social en el suroccidente de Guatemala?

Determinar cuáles son los factores físicos correspondientes a los hogares que son significativos en los indicadores nutricionales infantiles, y en qué niveles estos factores provocan valores nutricionales bajos o deficientes, para mejorar las propuestas de desarrollo social en el suroccidente de Guatemala, mediante un modelo de regresión logística multivariada y el análisis de odds ratios.

Los correspondientes a niveles asociados a la desnutrición infantil son:

Variable	Número de hogares	Porcentaje de hogares	Valor de Odds Ratio	Intervalo de Confianza
1. Agua	1	100%	1.000	0.999 - 1.001
2. Agua	1	100%	1.000	0.999 - 1.001
3. Agua	1	100%	1.000	0.999 - 1.001
4. Agua	1	100%	1.000	0.999 - 1.001
5. Agua	1	100%	1.000	0.999 - 1.001
6. Agua	1	100%	1.000	0.999 - 1.001
7. Agua	1	100%	1.000	0.999 - 1.001
8. Agua	1	100%	1.000	0.999 - 1.001
9. Agua	1	100%	1.000	0.999 - 1.001
10. Agua	1	100%	1.000	0.999 - 1.001

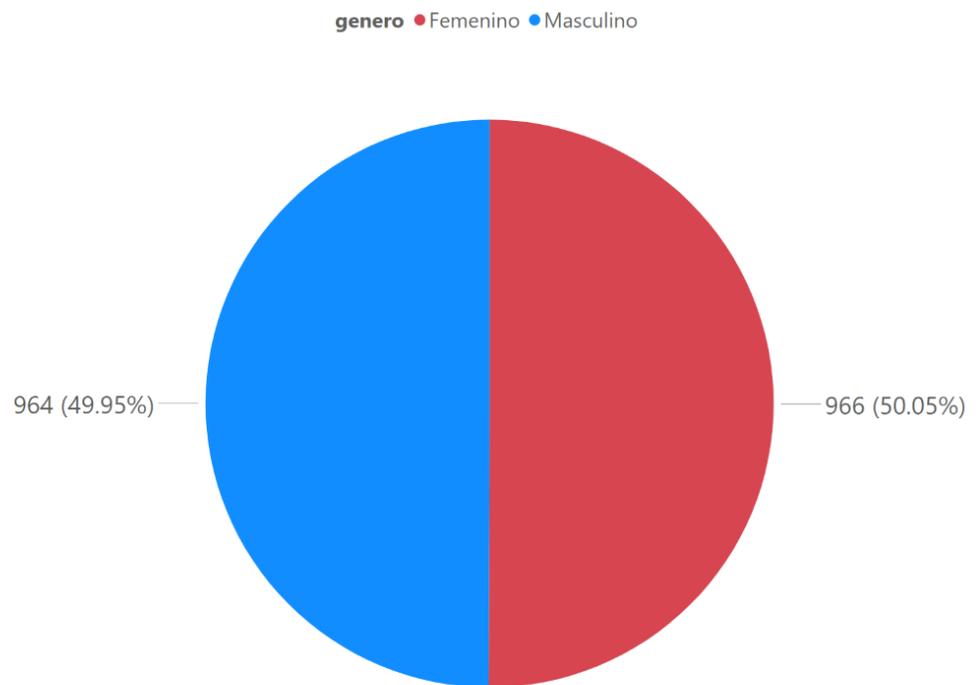
Se determinaron los factores físicos determinantes para explicar y pronosticar condiciones de desnutrición infantil utilizando el análisis de devianza y los odds ratios obtenidos con los modelos de regresión desarrollados. Los factores y niveles identificados fueron la fuente de agua cuando no es tubería dentro de vivienda, el piso de tierra de la casa y la disposición de excretas que no esté constituida por inodoro instalado.

Se hace necesario dar seguimiento a los factores explicativos determinados como significativos, proponiendo análisis específicos, con codificación modificada, segmentación de niveles adicionales, y otros cambios que se consideren necesarios.

Nota. Elaboración propia, realizado con Microsoft Word.

Apéndice 3.

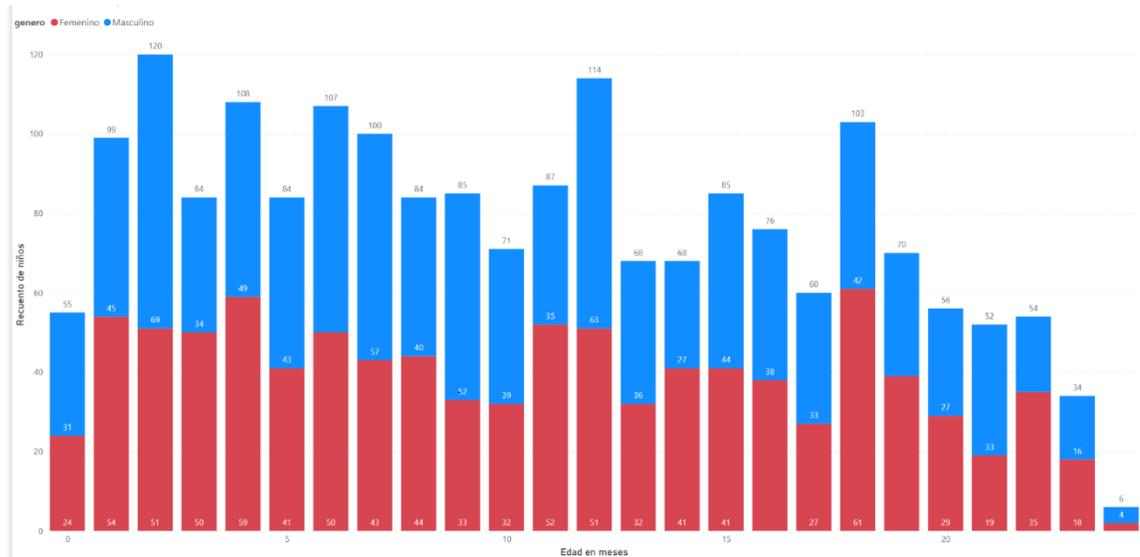
Muestra de niños por género



Nota. Elaboración propia, realizado con Microsoft Excel.

Apéndice 4.

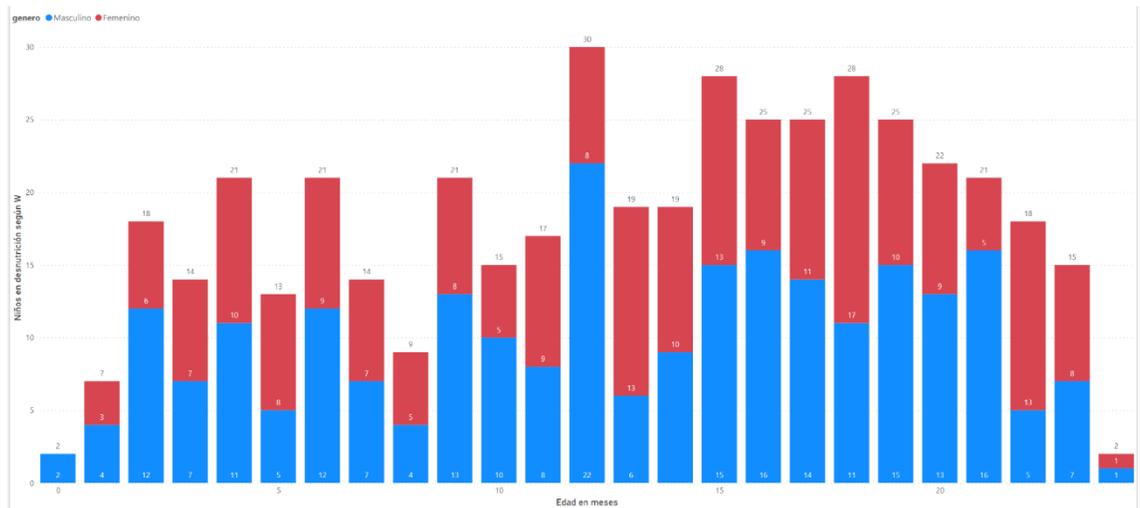
Muestra de niños por género y edad



Nota. Elaboración propia, realizado con Microsoft Excel.

Apéndice 5.

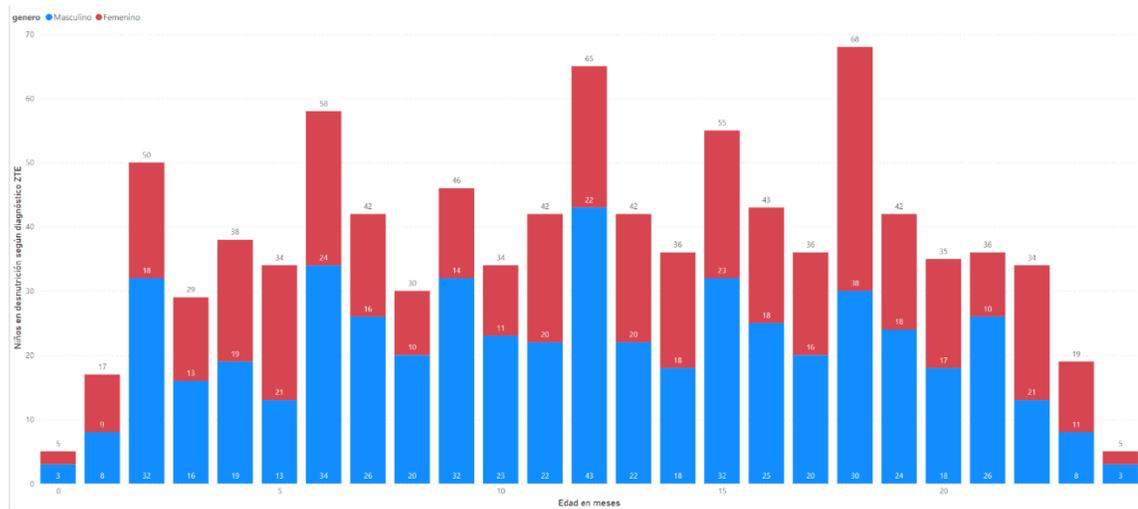
Diagnóstico Waterlow por género



Nota. Elaboración propia, realizado con Microsoft Excel.

Apéndice 6.

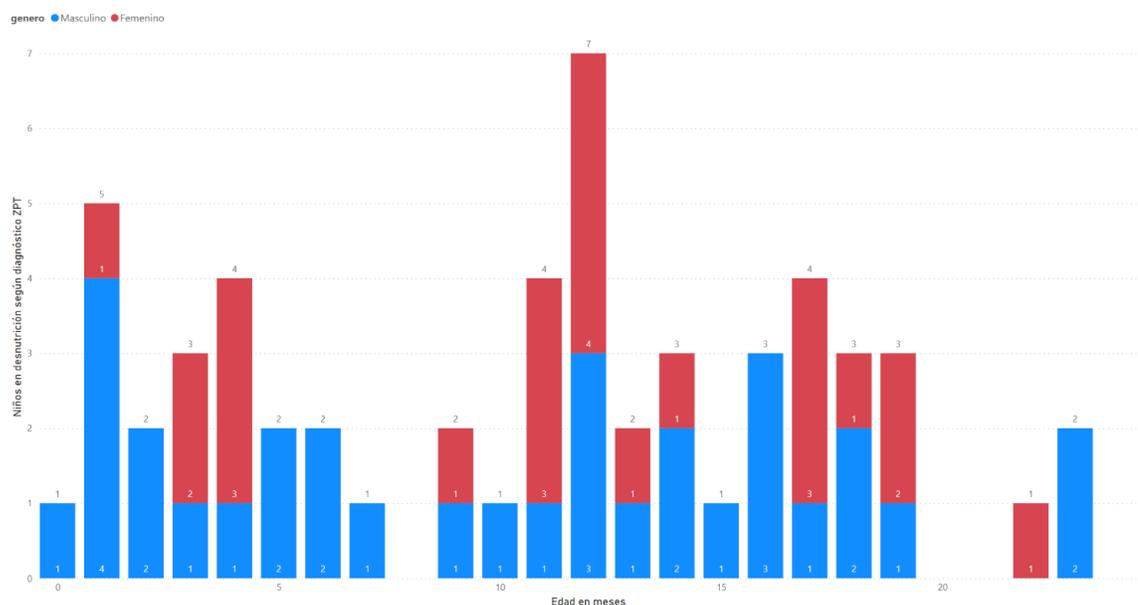
Diagnostico ZTE por género



Nota. Elaboración propia, realizado con Microsoft Excel.

Apéndice 7.

Diagnóstico ZPT por género



Nota. Elaboración propia, realizado con Microsoft Excel.

ANEXOS

Anexo 1.

Tabla talla para la edad en niños OMS

Simplified field tables

Length-for-age BOYS Birth to 2 years (percentiles)		 World Health Organization				
Year: Month	Months	3rd	15th	Median	85th	97th
0: 0	0	46.3	47.9	49.9	51.8	53.4
0: 1	1	51.1	52.7	54.7	56.7	58.4
0: 2	2	54.7	56.4	58.4	60.5	62.2
0: 3	3	57.6	59.3	61.4	63.5	65.3
0: 4	4	60.0	61.7	63.9	66.0	67.8
0: 5	5	61.9	63.7	65.9	68.1	69.9
0: 6	6	63.6	65.4	67.6	69.8	71.6
0: 7	7	65.1	66.9	69.2	71.4	73.2
0: 8	8	66.5	68.3	70.6	72.9	74.7
0: 9	9	67.7	69.6	72.0	74.3	76.2
0:10	10	69.0	70.9	73.3	75.6	77.6
0:11	11	70.2	72.1	74.5	77.0	78.9
1: 0	12	71.3	73.3	75.7	78.2	80.2
1: 1	13	72.4	74.4	76.9	79.4	81.5
1: 2	14	73.4	75.5	78.0	80.6	82.7
1: 3	15	74.4	76.5	79.1	81.8	83.9
1: 4	16	75.4	77.5	80.2	82.9	85.1
1: 5	17	76.3	78.5	81.2	84.0	86.2
1: 6	18	77.2	79.5	82.3	85.1	87.3
1: 7	19	78.1	80.4	83.2	86.1	88.4
1: 8	20	78.9	81.3	84.2	87.1	89.5
1: 9	21	79.7	82.2	85.1	88.1	90.5
1:10	22	80.5	83.0	86.0	89.1	91.6
1:11	23	81.3	83.8	86.9	90.0	92.6
2: 0	24	82.1	84.6	87.8	91.0	93.6

WHO Child Growth Standards

Nota. Obtenido de OMS (2017). *talla para la edad en niños OMS.* (https://www.aepap.org/sites/default/files/curvas_oms.pdf), consultado el 21 de septiembre de 2021. De dominio público.

Anexo 2.

Tabla talla para la edad en niñas OMS

Simplified field tables

Length-for-age GIRLS Birth to 2 years (percentiles)		 World Health Organization				
Year: Month	Months	3rd	15th	Median	85th	97th
0: 0	0	45.6	47.2	49.1	51.1	52.7
0: 1	1	50.0	51.7	53.7	55.7	57.4
0: 2	2	53.2	55.0	57.1	59.2	60.9
0: 3	3	55.8	57.6	59.8	62.0	63.8
0: 4	4	58.0	59.8	62.1	64.3	66.2
0: 5	5	59.9	61.7	64.0	66.3	68.2
0: 6	6	61.5	63.4	65.7	68.1	70.0
0: 7	7	62.9	64.9	67.3	69.7	71.6
0: 8	8	64.3	66.3	68.7	71.2	73.2
0: 9	9	65.6	67.6	70.1	72.6	74.7
0:10	10	66.8	68.9	71.5	74.0	76.1
0:11	11	68.0	70.2	72.8	75.4	77.5
1: 0	12	69.2	71.3	74.0	76.7	78.9
1: 1	13	70.3	72.5	75.2	77.9	80.2
1: 2	14	71.3	73.6	76.4	79.2	81.4
1: 3	15	72.4	74.7	77.5	80.3	82.7
1: 4	16	73.3	75.7	78.6	81.5	83.9
1: 5	17	74.3	76.7	79.7	82.6	85.0
1: 6	18	75.2	77.7	80.7	83.7	86.2
1: 7	19	76.2	78.7	81.7	84.8	87.3
1: 8	20	77.0	79.6	82.7	85.8	88.4
1: 9	21	77.9	80.5	83.7	86.8	89.4
1:10	22	78.7	81.4	84.6	87.8	90.5
1:11	23	79.6	82.2	85.5	88.8	91.5
2: 0	24	80.3	83.1	86.4	89.8	92.5

WHO Child Growth Standards

Nota. Obtenido de OMS (2017). *talla para la edad en niños OMS.* (https://www.aepap.org/sites/default/files/curvas_oms.pdf), consultado el 21 de septiembre de 2021. De dominio público.

Anexo 3.

Tabla peso para la talla en niños OMS

Simplified field tables

Weight-for-length BOYS Birth to 2 years (percentiles)		 World Health Organization			
cm	3rd	15th	Median	85th	97th
45.0	2.1	2.2	2.4	2.7	2.9
45.5	2.1	2.3	2.5	2.8	3.0
46.0	2.2	2.4	2.6	2.9	3.1
46.5	2.3	2.5	2.7	3.0	3.2
47.0	2.4	2.5	2.8	3.1	3.3
47.5	2.4	2.6	2.9	3.1	3.4
48.0	2.5	2.7	2.9	3.2	3.5
48.5	2.6	2.8	3.0	3.3	3.6
49.0	2.7	2.9	3.1	3.4	3.7
49.5	2.7	2.9	3.2	3.5	3.8
50.0	2.8	3.0	3.3	3.7	4.0
50.5	2.9	3.1	3.4	3.8	4.1
51.0	3.0	3.2	3.5	3.9	4.2
51.5	3.1	3.3	3.6	4.0	4.3
52.0	3.2	3.4	3.8	4.1	4.5
52.5	3.3	3.6	3.9	4.3	4.6
53.0	3.4	3.7	4.0	4.4	4.7
53.5	3.5	3.8	4.1	4.5	4.9
54.0	3.6	3.9	4.3	4.7	5.0
54.5	3.8	4.0	4.4	4.8	5.2
55.0	3.9	4.2	4.5	5.0	5.4
55.5	4.0	4.3	4.7	5.1	5.5
56.0	4.1	4.4	4.8	5.3	5.7
56.5	4.3	4.6	5.0	5.4	5.9
57.0	4.4	4.7	5.1	5.6	6.0
57.5	4.5	4.8	5.3	5.8	6.2
58.0	4.6	5.0	5.4	5.9	6.4
58.5	4.8	5.1	5.6	6.1	6.5
59.0	4.9	5.2	5.7	6.2	6.7
59.5	5.0	5.4	5.9	6.4	6.9

Nota. Obtenido de OMS (2017). *Peso para la talla en niños OMS.* (https://www.aepap.org/sites/default/files/curvas_oms.pdf), consultado el 21 de septiembre de 2021). De dominio público.

Anexo 4.

Tabla peso para la talla en niñas OMS

Simplified field tables

Weight-for-length GIRLS Birth to 2 years (percentiles)				World Health Organization	
cm	3rd	15th	Median	85th	97th
45.0	2.1	2.2	2.5	2.7	2.9
45.5	2.2	2.3	2.5	2.8	3.0
46.0	2.2	2.4	2.6	2.9	3.1
46.5	2.3	2.5	2.7	3.0	3.2
47.0	2.4	2.6	2.8	3.1	3.3
47.5	2.4	2.6	2.9	3.2	3.4
48.0	2.5	2.7	3.0	3.3	3.5
48.5	2.6	2.8	3.1	3.4	3.7
49.0	2.7	2.9	3.2	3.5	3.8
49.5	2.8	3.0	3.3	3.6	3.9
50.0	2.8	3.1	3.4	3.7	4.0
50.5	2.9	3.2	3.5	3.8	4.1
51.0	3.0	3.2	3.6	3.9	4.3
51.5	3.1	3.4	3.7	4.0	4.4
52.0	3.2	3.5	3.8	4.2	4.5
52.5	3.3	3.6	3.9	4.3	4.7
53.0	3.4	3.7	4.0	4.4	4.8
53.5	3.5	3.8	4.2	4.6	5.0
54.0	3.6	3.9	4.3	4.7	5.1
54.5	3.7	4.0	4.4	4.9	5.3
55.0	3.9	4.1	4.5	5.0	5.4
55.5	4.0	4.3	4.7	5.2	5.6
56.0	4.1	4.4	4.8	5.3	5.8
56.5	4.2	4.5	5.0	5.5	5.9
57.0	4.3	4.6	5.1	5.6	6.1
57.5	4.4	4.8	5.2	5.7	6.2
58.0	4.5	4.9	5.4	5.9	6.4
58.5	4.6	5.0	5.5	6.0	6.5
59.0	4.8	5.1	5.6	6.2	6.7
59.5	4.9	5.2	5.7	6.3	6.9

Nota. Obtenido de OMS (2017). *Peso para la talla en niñas OMS.* (https://www.aepap.org/sites/default/files/curvas_oms.pdf), consultado el 21 de septiembre de 2021). De dominio público.