



UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE ESTADÍSTICA
MAGÍSTER ESTADÍSTICA



**FACTORES QUE EXPLICAN EL SOBREPESO Y LA
OBESIDAD EN CHILE.**

Tesis presentada a la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Concepción para optar al grado académico de Magíster en Estadística.

Por: Francisca Romo Muñoz
Profesor Guía: Sebastián Niklitschek Soto
Profesor Co-Guía: Rodrigo Romo Muñoz

2022
Concepción, Chile



UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE ESTADÍSTICA
MAGÍSTER ESTADÍSTICA



**FACTORES QUE EXPLICAN EL SOBREPESO Y LA
OBESIDAD EN CHILE.**

Profesor Guía : Sebastián Niklitschek Soto Firma

Profesor Co-Guía : Rodrigo Romo Muñoz Firma

Profesor Colaborador : Joel Muñoz Gutiérrez Firma

Profesor Consejero : Katia Sáez Carrillo Firma

Nombre Tesista : Francisca Romo Muñoz Firma

2022

Concepción, Chile



Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento.



*Dedicada a
mi familia.*

Agradecimientos

Mi más profundo agradecimiento a mi hermano Rodrigo, por guiarme en cada momento de mi vida y siempre confiar en mis capacidades, por cada consejo entregado, por la atención dedicada y por el apoyo incondicional. Le agradezco a mi hermano Daniel, por confiar en mis conocimientos y estar siempre dispuesto a ayudarme.

A mis padres, Rosa y Ricardo les agradezco por darme la oportunidad de estudiar en la Universidad de Concepción, por guiarme a través del camino de la vida, por apoyarme en cada decisión, por creer en mí.

No puedo dejar de agradecer a Diego, que cada día tuvo una palabra de aliento para animarme, por su comprensión y cariño entregado cada día.

A los profesores del Departamento de Estadística, gracias por los conocimientos brindados, por el tiempo dedicado a la enseñanza, por su preocupación y buena disposición siempre. A mis amigos y compañeros de universidad, gracias por cada momento compartido en estos años.

A Dios por iluminarme en cada momento. Gracias a la vida por este nuevo logro, gracias a cada una de las personas que me apoyaron y creyeron en la realización de esta investigación.

Resumen

El sobrepeso y la obesidad son considerados actualmente como el mayor problema de salud pública en Chile. En el año 2017, el 74,2% de la población chilena presentó exceso de peso. En esta investigación se analiza distintos factores en la explicación del sobrepeso y obesidad, entre ellos factores sociodemográficos, conductuales y ambientales, y nuevos factores propuestos por la literatura como factor psicosocial, enfermedades diagnosticadas, exámenes de laboratorio y medidas biofisiológicas.

Se utilizó la III Encuesta Nacional de Salud (ENS) aplicada el año 2017 por el Ministerio de Salud, corresponde a una encuesta poblacional de tipo transversal, con un total de 6233 encuestados y 5520 exámenes de laboratorio. Un modelo de regresión logística multinomial fue utilizado para la determinación de las relaciones existentes entre el sobrepeso y obesidad con los distintos factores considerados. Características como sexo, estado civil, zona geográfica, consumo de alcohol y cigarrillos, estrés, estrés financiero, enfermedades diagnosticadas como diabetes e hipertensión, y algunos resultados a exámenes médicos son importantes en la determinación del estado nutricional.

Los resultados que arroja esta investigación, en un futuro serán importantes a la hora de definir políticas públicas enfocadas en el área de salud, particularmente en intervenciones contra el sobrepeso y obesidad en Chile.

Índice general

Resumen	II
1. Introducción	1
2. Marco Teórico	6
2.1. Sobrepeso y Obesidad	6
2.2. Encuesta Nacional de Salud	11
3. Materiales y Métodos	15
3.1. III Encuesta Nacional de Salud	15
3.2. Técnicas Estadísticas	21
3.2.1. Árboles de Decisión	23
3.2.2. Bosque Aleatorio	26
3.2.3. Regresión Logística Multinomial	29
4. Resultados	32
4.1. Análisis descriptivo	32
4.2. Selección de variables	37
4.3. Regresión Logística Multinomial	39
5. Conclusión	49
Anexos	
A. Descripción módulos	53
B. Resumen gráfico de variables	56
Referencias	63

Índice de cuadros

2.1.1.	Clasificación Índice de Masa Corporal.	7
2.2.1.	Comparación resultados ENS 2003-2010.	13
2.2.2.	Comparación resultados ENS 2003-2010 por sexo.	13
3.1.1.	Ficha técnica de la muestra.	16
3.1.2.	Variables utilizadas en los modelos y su respectiva fuente en los formularios de la ENS 2016-2017.	19
4.1.1.	Caracterización de la muestra basada en el estado nutricional.	32
4.2.1.	Resultados entrenamiento bosques aleatorios.	37
4.2.2.	Importancia de variables a través de bosque aleatorio.	38
4.3.1.	Variables utilizadas en estimación del modelo logístico multinomial para el estado nutricional.	40
4.3.2.	Resultados estimación del modelo logístico multinomial para el estado nutricional.	42
4.3.3.	Tabla de clasificación del modelo logístico multinomial para el estado nutricional.	48
A.0.1.	Descripción módulos formulario 1 y 2 de la III Encuesta Nacional de Salud.	53

Índice de figuras

3.2.1.	Diagrama árbol de decisión.	24
3.2.2.	Ejemplo nodo padre y nodo hijo.	25
B.0.1.	Porcentaje del estado nutricional dentro de la variable sexo. . .	56
B.0.2.	Porcentaje del estado nutricional dentro de la variable estado civil. . .	57
B.0.3.	Porcentaje del estado nutricional dentro de la variable sexo x estado civil.	57
B.0.4.	Porcentaje del estado nutricional dentro de la variable edad. . .	58
B.0.5.	Porcentaje del estado nutricional dentro de la variable zona geográfica.	58
B.0.6.	Porcentaje del estado nutricional dentro de la variable área geográfica.	59
B.0.7.	Porcentaje del estado nutricional dentro de la variable nivel educacional.	59
B.0.8.	Porcentaje del estado nutricional dentro de la variable consumo de alcohol.	60
B.0.9.	Porcentaje del estado nutricional dentro de la variable consumo de cigarrillos.	60
B.0.10.	Porcentaje del estado nutricional dentro de la variable dieta saludable.	61
B.0.11.	Porcentaje del estado nutricional dentro de la variable actividad física.	61
B.0.12.	Porcentaje del estado nutricional dentro de la variable estrés. . .	62

Capítulo 1

Introducción

En el año 2006, la Organización Mundial de la Salud señaló que el sobrepeso y la obesidad alcanzaron caracteres de epidemia mundial, en donde más de 1000 millones de personas adultas presentan exceso de peso, y al menos 300 millones de ellas son obesas (Ministerio de Salud [MINSAL], 2006)

La prevalencia del sobrepeso y obesidad ha ido en aumento en todo el mundo, y por esto se ha convertido en un importante problema de salud pública que genera un alto gasto sanitario. En Chile las cifras del sector de la salud son:

- i) Según lo informado por la Organización para la Cooperación y Desarrollo Económico (OECD por sus siglas en inglés), para el año 2018 el 5,2% del PIB se destinó al sector de la salud, cifra que está por debajo del 8% del gasto promedio de los países miembros de la organización (Organization for Economic Co-operation and Development [OECD], 2019a).
- ii) El 20,2% del presupuesto fiscal del año 2019 se destinó al Ministerio de Salud, según Cerda (2018).

Pese al menor gasto en salud que realiza Chile en comparación al promedio de los países OECD, el país tiene un sistema de salud en buen funcionamiento,

bien organizado y gobernado de forma efectiva (Organization for Economic Co-operation and Development [OECD], 2019b). Para los sistemas de salud el objetivo principal es abarcar todas las actividades que tienen por finalidad promover, restablecer o mantener la salud (World Health Organization [WHO], 2000). Previo a la pandemia ocasionada por el coronavirus (COVID-19), originada por el síndrome respiratorio agudo grave (SARS-CoV-2) en la ciudad de Wuhan, capital de la provincia de Hubei en China central, los sistemas de salud mundial se enfrentaban a la pandemia del sobrepeso y obesidad, ya que para el año 2015 se presentaron 107,7 millones de niños y 603,7 millones de adultos obesos (Afshin *et al.*, 2017). En el año 2016, más de 1900 millones de adultos de 18 o más años tenían sobrepeso, de los cuales, más de 650 millones eran obesos; esta última cifra representa alrededor del 13% de la población adulta mundial (World Health Organization [WHO], 2021).

Actualmente, el mayor desafío para el sistema sanitario es reducir el sobrepeso y la obesidad (OECD, 2019b), ya que según lo informado por el reporte “Encuesta Nacional de Salud 2016-2017. Primeros resultados” de Departamento de Epidemiología [EPI] (2021), el 74,2% de la población en Chile presenta sobrepeso u obesidad (39,8% y 34,4% respectivamente). La Junta Nacional de Auxilio Escolar y Becas ha reportado que en menores a 15 años, se presenta un 52% con sobrepeso u obesidad (Junta Nacional de Auxilio Escolar y Becas [JUNAEB], 2019), lo cual, junto con el resultado anterior, posiciona a Chile como uno de los países con las tasas de sobrepeso y obesidad más altas de todos los países pertenecientes a la OECD, solo superado por Estado Unidos (Organization for Economic Co-operation and Development [OECD], 2019c).

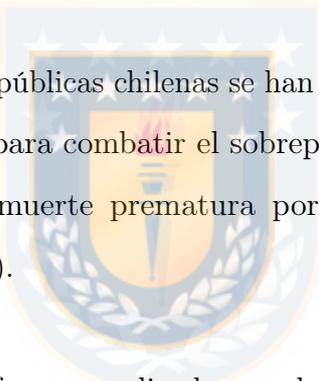
El sobrepeso y la obesidad no son solo un problema estético que afecta al bienestar de las personas, sino que también están relacionados con una serie de

consecuencias para la salud, ya sea cuando se presenta en la infancia o en la vida adulta, de forma transversal en todos los países. La presencia de sobrepeso u obesidad en la niñez está fuertemente asociada con la obesidad adulta, y hay evidencia que muestra que ésta se puede mantener en el tiempo (Singh *et al.*, 2008; Llewellyn *et al.*, 2016). Además, los niños con sobrepeso u obesidad presentan problemas psicológicos o psiquiátricos, como la baja autoestima o problemas de comportamiento, y problemas de salud, entre ellos problemas cardiovasculares (presión arterial alta, dislipidemia, anomalías en la función del ventrículo izquierdo; anomalías en la función endotelial y resistencia a la insulina), asma, y riesgo de desarrollar diabetes tipo I (Reilly *et al.*, 2003).

Por otro lado, el sobrepeso y obesidad en adultos presenta efectos adversos en los logros sociales y económicos (nivel de ingreso y logro educativo) (Reilly *et al.*, 2003), y es considerada un factor de riesgo significativamente mayor para una serie de enfermedades. Entre estas se incluyen las enfermedades cardiovasculares o ECV (cardiopatía isquémica, accidente cerebro-vascular isquémico, accidente cerebro-vascular hemorrágico, enfermedad cardíaca hipertensiva), la diabetes tipo II y distintos tipos de cáncer (de colon y recto, esófago, hígado, mama, vesícula, útero, ovarios, riñón, tiroides y leucemia), lo cual genera una mortalidad prematura (Afshin *et al.*, 2017; Llewellyn *et al.*, 2016).

El sobrepeso y obesidad es el resultado de una combinación de factores como factor genético, conductual y factor ambiental (Xu y Xue, 2016). La promoción de la salud en Chile debe ser completa, considerando la prevención del consumo de tabaco y alcohol, fomentando la alimentación saludable, actividad física y salud mental. Es en la alimentación saludable donde son importantes las variables socioeconómicas, ya que del ingreso total por hogar va a depender el tipo de alimentación y opciones de realizar actividad física (Bennett *et al.*, 2008).

Es importante mencionar que algunas condiciones médicas son causas del sobrepeso y obesidad, síndromes genéticos como Prader-Willi, Bardet-Biedl o Cohen están asociados con el sobrepeso y la obesidad, al igual que algunos trastornos endocrinos como el hipotiroidismo o tumores (National Heart Lung and Blood Institute [NIH], 2021). Por lo tanto, los problemas de salud y resultados de exámenes médicos son importantes a la hora de identificar factores asociados al sobrepeso y obesidad. El estrés, la depresión e indicadores del bienestar emocional forman parte del factor psicosocial, y se han presentado numerosos estudios que indican una fuerte asociación con el sobrepeso y obesidad (Gundersen *et al.*, 2011).



Actualmente las políticas públicas chilenas se han enfocado solo en los factores conductuales y ambientales para combatir el sobrepeso y la obesidad, y de esta forma reducir el riesgo de muerte prematura por enfermedades crónicas no transmisibles (OECD, 2019b).

A pesar de los grandes esfuerzos realizados por las instituciones nacionales de Chile, el sobrepeso y la obesidad han seguido aumentando, existen investigaciones que analizan el sobrepeso y obesidad considerando un factor o algunos de los mencionados anteriormente (Afshin *et al.*, 2017). Es por esta razón que resulta prioritario realizar investigaciones que permitan conocer los principales determinantes para luchar contra el sobrepeso y la obesidad considerando todos los factores anteriores. El objetivo de esta investigación es modelar el sobrepeso y la obesidad en Chile a través de variables sociodemográficas y distintos factores, utilizando la III Encuesta Nacional de Salud aplicada en el año 2017, adicionalmente se incluye el factor psicosocial, enfermedades y exámenes de laboratorio.

Para tales efectos, esta tesis se organiza en 5 capítulos, siendo el primer capítulo la introducción y motivación a la investigación. Para el segundo capítulo, se presenta un resumen que contiene las principales referencias bibliográficas disponibles para estudiar el problema del sobrepeso y obesidad en Chile. Luego en el tercer capítulo se hace referencia a la procedencia de los datos y las herramientas estadísticas para realizar la investigación, el siguiente capítulo presenta los resultados obtenidos del análisis propuesto. Finalmente, el último capítulo presenta conclusiones a partir de los resultados obtenidos. Para finalizar, se presentan las distintas referencias bibliográficas utilizadas a lo largo de la investigación.



Capítulo 2

Marco Teórico

2.1. Sobrepeso y Obesidad

El sobrepeso y la obesidad son definidos como una enfermedad crónica no transmisible (ECNT, denominación de la Organización Mundial de la Salud), que se debe a una acumulación anormal o excesiva de grasa que puede ser perjudicial para la salud, dentro de estas enfermedades también se encuentran la diabetes, enfermedades cardiovasculares, cáncer, osteoporosis, malformaciones congénitas y enfermedades del tipo inmune, que son patologías que están directa o indirectamente relacionadas con la alimentación y por lo tanto se pudieran prevenir (MINSAL, 2006). Por otra parte, también se define al sobrepeso y la obesidad como “un importante problema de salud pública en la mayoría de los países, que conlleva a la disminución de la esperanza de vida y a generar altos costos sociales y económicos” (Atalah, 2012).

Una forma de medir el sobrepeso y la obesidad es mediante el cálculo del Índice de Masa Corporal (IMC), esto es el peso de una persona en kilogramos dividido por el cuadrado de su estatura en metros. El IMC clasifica a las personas como bajo peso, normal, sobrepeso y obesidad en 3 clases (Cuadro 2.1.1), para este trabajo se consideró obesidad cuando el IMC es mayor o igual a 30, sin

importar la clase. A continuación, se muestra la clasificación del estado nutricional según el índice de masa corporal:

Cuadro 2.1.1: Clasificación Índice de Masa Corporal.

Estado Nutricional	IMC
Bajo peso	$< 18,5$
Peso normal	$18,5 - 24,9$
Sobrepeso	$25,0 - 29,9$
Obesidad Clase I	$30,0 - 34,9$
Obesidad Clase II	$35,0 - 39,9$
Obesidad Clase III	≥ 40

Fuente: World Health Organization [WHO], 2020.

Esta enfermedad se ha incrementado de forma alarmante en los países con alto desarrollo humano, llegando a considerarse como el problema de salud pública mayor del siglo XXI (Castro *et al.*, 2018); en nuestro país los datos también son alarmantes, Schubert y Ávalos (2020) señala que el 67% de la población muestra algún grado de sobrepeso u obesidad, con un 39,9% que posee sobrepeso y un 27% posee obesidad.

El sobrepeso y la obesidad pueden llegar a reducir la esperanza de vida de una persona hasta en diez años, es por las consecuencias y cifras alarmantes que hacen de esta enfermedad uno de los mayores retos de la salud pública para el siglo XXI (Ministerio de Salud [MINSAL], 2007). Actualmente el sistema de salud en Chile se ve enfrentado a importantes desafíos, como disminuir la tasa de personas fumadoras, disminuir el consumo de alcohol, controlar la mortalidad por cáncer y reducir el sobrepeso y la obesidad. Debido a los distintos desafíos y la alta preocupación por el sistema de salud, se crea la Encuesta Nacional de Salud (ENS) que es utilizada por el Ministerio de Salud para saber qué enfermedades y tratamientos están recibiendo las personas adultas que residen en el país.

Por otra parte, frente al acelerado incremento del sobrepeso y obesidad en Chile, se efectuaron diferentes programas para enfrentar la problemática desde la alimentación y actividad física, como: VIDA Chile (entre los años 1998 y 2005); la Estrategia Global contra la Obesidad (entre los años 2006 y 2010); y el programa Elige Vivir Sano (entre los años 2011 y 2014); pero ninguno de estos han sido suficientes para cambiar la tendencia al alza del sobrepeso y obesidad Vio *et al.* (2008).

Dado que en los últimos años no se han visto acciones exitosas en el control del sobrepeso y la obesidad, Vio del Rio (2018) comenta que existe un consenso en que las políticas públicas en Chile deben seguir actuando en tres niveles:

- i) En primer lugar, enfocarse en el entorno escolar, entregando programas de educación nutricional, normas de alimentación de frutas y verduras para la buena nutrición y alimentación de los menores.
- ii) En segundo lugar, continuar con los impuestos a los alimentos procesados que contienen alto nivel de grasas, azúcar y sal, a bebidas azucaradas, junto con subsidiar los alimentos saludables.
- iii) Por último, focalizarse en el etiquetado nutricional de los productos alimenticios, centrado en los alimentos altos en calorías, grasas saturadas, azúcares y sodio.

Lo mencionado anteriormente considera que la promoción de la vida saludable debe ser abordada de forma integral, es decir, considerar prevención del consumo de tabaco y alcohol, educar y fomentar la alimentación saludable, la actividad física y la salud mental, y también contar con el apoyo y colaboración de la sociedad. Frente a esto, Vio del Rio (2018) indica que la educación de la población en temas de alimentación y estilos de vida saludables, desde las edades más tempranas, pasa a ser un tema crucial para el futuro de la salud pública en Chile y en el mundo.

Sin embargo, las condiciones esperadas mencionadas anteriormente, se contraponen a la dinámica de la sociedad, ya que Chile está inmerso en un contexto económico, social y cultural que contribuye a la práctica de estilos de vida muy poco saludables. Los cambios demográficos que actualmente está viviendo el país, como el avance tecnológico, la desigualdad social, y el gran desconocimiento sobre los efectos nocivos del sedentarismo, se ha traducido en que más de un 80 % de la población adulta del país sea sedentaria, según datos de la Encuesta Nacional de Salud año 2010 (Departamento de Epidemiología [EPI], 2017).

Según Xu y Xue (2016), el sobrepeso y la obesidad son generadas por causas multidimensionales como:

- i) Factores genéticos
- ii) Factores conductuales
- iii) Factores ambientales



El factor genético, que se transmite a través de los genes de los padres, ha sido el menos abordado por las políticas de salud pública de Chile. El factor conductual está asociado a la adopción de estilos de vida saludables por parte de la población, entre los que se encuentran tener una dieta balanceada, evitar consumo de bebidas azucaradas, realizar actividad física, evitar fumar tabaco y prevenir el consumo de alcohol, con excepción del consumo moderado de vino tinto (Romo-Muñoz *et al.*, 2020). El seguir una dieta balanceada se considera que es lo más fácil de adoptar y es el factor más estudiado en la obesidad, debido a que en la dieta de la población chilena se destaca el consumo de alimentos ultraprocesados, lo cual ha generado una de las mayores tasas de obesidad en el mundo (Schubert y Ávalos, 2020). Por último, el factor ambiental se refiere a los entornos en donde interactúa una persona, en el caso de los niños, estos son el

hogar, la escuela y los entornos comunitarios (Xu y Xue, 2016), y para los adultos se considera el lugar de trabajo y el hogar.

Precisamente, son los factores conductuales y ambientales en los que se han enfocado las políticas públicas chilenas para combatir el sobrepeso y la obesidad, y las enfermedades crónicas que causan y, de esta forma, reducir el riesgo de muerte prematura por enfermedades crónicas no transmisibles. En este sentido, Chile ha implementado una amplia gama de políticas e intervenciones para combatir el sobrepeso y la obesidad, que consideran intervenir específicamente las dietas poco saludables, impactando tanto el ambiente escolar de los niños y adolescentes, como la actividad física (OECD, 2019b). Estos paquetes de políticas presentan diferentes alcances. Algunos se han generado implementando leyes a nivel nacional que solo consideran las dietas (campañas en los medios de comunicación, etiquetado de alimentos, restricciones de la publicidad, impuesto al azúcar); también intervenciones a nivel regional que consideran dietas y actividad física (programas intersectoriales de salud de la población; la reforma de los espacios públicos); otras consisten en programas en instituciones públicas y privadas (intervenciones en las escuelas; intervenciones en el lugar de trabajo); y, finalmente existen programas a nivel individual como consejería y clases de actividad en la atención primaria (OECD, 2019b).

Adicional a los factores mencionados, el sobrepeso y la obesidad como problemas de salud también presentan componentes relacionados al ámbito social, económico y político que influyen en los niveles de exposición y vulnerabilidad de la población (EPI, 2017).

2.2. Encuesta Nacional de Salud

En Chile desde el año 2000, el Ministerio de Salud ha realizado esfuerzos para obtener información sobre las enfermedades no transmisibles y sus principales factores de riesgos. Para lograrlo, se ha sumado a las recomendaciones internacionales, desarrollando encuestas poblacionales bajo el marco de un método progresivo. En el año 2000 y 2006 se realizó el cuestionario de autoreporte denominado Encuesta de Calidad de Vida. Posteriormente a este cuestionario surge la Encuesta Nacional de Salud (ENS), la que incorporó además mediciones biofisiológicas y exámenes de laboratorio de la población (Ministerio de Salud [MINSAL], 2011).

La Encuesta Nacional de Salud es una herramienta para saber qué enfermedades tiene la población mayor a 14 años que reside en Chile. A partir de sus resultados se pretende estimar la prevalencia de problemas de salud prioritarios en la población adulta chilena. Esto se logra a través de cuestionarios, mediciones biofisiológicas y exámenes de laboratorio. Los resultados entregados por la encuesta permiten obtener conocimiento sólido para realizar una planificación sanitaria nacional y, a su vez hacer promoción y prevención de salud para mejorar el nivel de salud y calidad de vida de los chilenos (EPI, 2021).

La primera Encuesta Nacional de Salud se realizó en el año 2003, y constituyó una fuente de información de alta calidad para la planificación sanitaria, para evaluar el estado de avance de los objetivos sanitarios entre los años 2000-2010. La muestra lograda fue de 3619 personas mayores de 17 años residentes en Chile, 1647 hombres y 1972 mujeres, en esa ocasión se incluyeron 23 condiciones de salud y 12 exámenes de laboratorio (EPI, 2021).

Entre los años 2009 y 2010, se elaboró y aplicó una segunda versión de la ENS, esto con el objetivo de mantener una periodicidad en el estudio de enfermedades crónicas y sus determinantes. En esa oportunidad la muestra lograda fue de 5416 personas, con 2200 hombres y 3216 mujeres, todos mayores de 14 años residentes en Chile. Se incorporaron nuevas condiciones de salud y exámenes de laboratorios, logrando 42 y 22 respectivamente (EPI, 2021). En ambas oportunidades fue necesario un encuestador y un profesional de la salud para las mediciones biofisiológicas y exámenes de laboratorio.

La información que arroja la ENS 2003 ayudó a establecer los objetivos sanitarios para los siguientes años, uno de ellos fue la realización de estrategias de intervención en salud y la Estrategia Global contra la Obesidad (EGO CHILE). La meta principal de EGO CHILE, es disminuir la prevalencia del sobrepeso y obesidad en nuestro país, fomentando una alimentación saludable y promoviendo la actividad física, a lo largo de toda la vida (MINSAL, 2006).

Dentro de las principales líneas de acción que se proponen para ser desarrolladas en el EGO Chile se encuentran (MINSAL, 2006):

- Fomentar políticas y planes de acción destinados a mejorar los hábitos alimentarios y aumentar la actividad física en la población.
- Sensibilizar e informar a la población del impacto positivo que, para su salud, tienen una alimentación equilibrada y la práctica regular de actividad física.
- Incentivar a las organizaciones comunitarias y redes sociales en la promoción de estilos de vida saludables.
- Incorporar en los programas educacionales actividades de promoción de estilos de vida saludables.

Luego de las estrategias y políticas de salud establecidas a partir de la información de la ENS 2003, es importante realizar una comparación con los resultados de la ENS 2010, para evidenciar los cambios en el estado nutricional y la prevalencia de los problemas de salud incluidos en ambas encuestas. A continuación se observan cuadros comparativos de resultados ENS 2003 y ENS 2010:

Cuadro 2.2.1: Comparación resultados ENS 2003-2010.

Problema de salud	Prevalencia	
	ENS 2003	ENS 2010
Sobrepeso	37,8 %	39,3 %
Obesidad	23,2 %	25,1 %
Tabaquismo	42,0 %	40,6 %
Hipertensión arterial	33,7 %	26,9 %
Diabetes	6,3 %	9,4 %
Sedentarismo	89,4 %	88,6 %
Síntomas depresivos en el último año	17,5 %	17,2 %

Fuente: Encuesta Nacional de Salud 2010 (MINSAL, 2011).

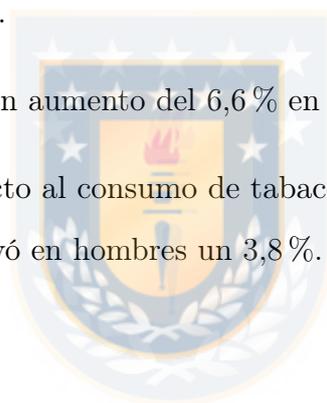
Cuadro 2.2.2: Comparación resultados ENS 2003-2010 por sexo.

Problema de salud	Prevalencia ENS 2003		Prevalencia ENS 2010		
	Hombres	Mujeres	Hombres	Mujeres	
Estado Nutricional	Bajo Peso	0,6 %	1,1 %	1,1 %	2,4 %
	Sobrepeso	43,2 %	32,7 %	45,3 %	33,6 %
	Obesidad	19,0 %	25,0 %	19,2 %	30,7 %
	Obesidad mórbida	0,2 %	2,3 %	1,3 %	3,3 %
Hipertensión	36,7 %	30,8 %	28,7 %	25,3 %	
Diabetes	4,8 %	3,8 %	8,4 %	10,4 %	
Dislipidemia	35,1 %	35,6 %	39,0 %	38,1 %	
Tabaquismo	48,0 %	37,0 %	44,2 %	37,1 %	
Sedentarismo	87,9 %	90,8 %	84,0 %	92,9 %	
Síndrome metabólico	23,0 %	22,3 %	41,7 %	31,0 %	
Síntomas depresivos en el último año	10,4 %	24,3 %	8,5 %	25,7 %	

Fuente: Resumen ejecutivo 2003 y 2009-2010 (EPI, 2021).

A partir de la información, se concluye:

- Del 2003 al 2010 la prevalencia del sobrepeso y obesidad aumentó, en 1,5 % y 1,9 % respectivamente.
- La hipertensión arterial y el sedentarismo, presentan descenso en la prevalencia total.
- Se presenta mayor prevalencia de obesidad, obesidad mórbida, sedentarismo y síntomas depresivos en el último año para mujeres en ambas encuestas.
- Los hombres tienen mayor prevalencia del sobrepeso y se mantiene para el ENS 2003 y ENS 2010.
- La diabetes presentó un aumento del 6,6 % en la prevalencia de mujeres.
- Por último, con respecto al consumo de tabaco se mantuvo la prevalencia en mujeres, y disminuyó en hombres un 3,8 %.



Capítulo 3

Materiales y Métodos

3.1. III Encuesta Nacional de Salud

La III Encuesta Nacional de Salud se aplicó entre agosto de 2016 y marzo de 2017, por un encuestador y un profesional de salud, según el tipo de cuestionario, a personas de 15 años o más, de nacionalidad chilena o extranjeros, que residen habitualmente en viviendas particulares ocupadas, localizadas en zonas urbanas y rurales de las quince regiones de Chile. La encuesta consta de 576 preguntas distribuidas en cuatro formularios, también se consideran mediciones antropométricas como peso, talla, circunferencia abdominal, circunferencia de brazo y pierna, adicionalmente se realizan distintos exámenes de sangre y orina (EPI, 2017). La ficha técnica de la muestra se encuentra en el Cuadro 3.1.1.

Cuadro 3.1.1: Ficha técnica de la muestra.

Organismo responsable	Departamento de Epidemiología MINSAL.
Diseño	Encuesta poblacional de tipo transversal.
Población objetivo	Personas de 15 años y más, chilenas o extranjeras que residen habitualmente en viviendas particulares ocupadas, localizadas en zonas urbanas o rurales de las quince regiones de Chile.
Representatividad	Nacional, regional y urbano/rural.
Modo de aplicación	Entrevista personal en hogar utilizando Tablet, aplicada por un encuestador y profesional enfermera considerando el tipo de cuestionario.
Período de trabajo de campo	Agosto de 2016 a marzo de 2017.
Tamaño muestral	6233 encuestados, de los cuales 5520 cuentan con exámenes de laboratorio de acuerdo a protocolo.
Error muestral	Error absoluto de muestreo de 2,6% a nivel nacional, raíz del efecto de diseño de 1,797, estimaciones con 95% de confianza y error relativo inferior a 30%.

Fuente: III ENS 2016-2017, primeros resultados (EPI, 2017).

La información se recolectó a partir de un muestreo de áreas geográficas en las 15 regiones del país, considerando zonas urbanas y rurales. El Departamento de Epidemiología (EPI) indica que se puede caracterizar el diseño muestral como probabilístico, estratificado geográficamente y multietápico, debido a que considera el diseño de una muestra principal, la muestra “ENS”, y tres muestras secundarias que considera a la submuestra de salud mental (CIDI), la de exámenes aleatorios (SMA), y la submuestra de tiroides (SMA2). Cabe destacar que también se considera una encuesta de desarrollo infantil (EDI) en todas las viviendas con niños entre 7 y 59 meses de edad.

La metodología para la selección de la muestra, según el Informe del Diseño Muestral de EPI (2016), comienza con la formación de 30 estratos de muestreo dados por el cruce entre región (15 regiones) y área (urbano y rural). Luego, para cada estrato se seleccionan comunas en la primera etapa de muestreo (las comunas con más de 30.000 habitantes en las áreas urbanas se incluyen con seguridad en la muestra). Para la segunda etapa de muestreo se seleccionan manzanas censales o localidades, para las áreas urbanas y rurales respectivamente. Por último, en la tercera etapa de muestreo se seleccionan las viviendas y, en la cuarta etapa de muestreo, se selecciona una persona de 15 o más años por vivienda.

A continuación se describen los formularios, según lo expuesto en el Informe de Diseño Muestral de EPI, 2021:

- **Formulario F1:** cuestionario aplicado por el encuestador en la primera visita a todas las personas seleccionadas en la muestra principal ENS.
- **Formulario F2:** cuestionario aplicado por una enfermera en la segunda visita, más algunas mediciones biofisiológicas realizadas también por la enfermera en dicha oportunidad.
- **Exámenes EX1:** corresponden a las muestras de sangre y orina recolectadas en la visita de la enfermera a todas las personas seleccionadas en la muestra principal ENS.
- **Formulario F4:** cuestionario aplicado por el encuestador en una tercera visita en todas las viviendas seleccionadas en la muestra principal ENS que tengan niños entre 7 y 59 meses de edad. Se aplican cuestionarios individuales por cada niño entre 7 y 59 meses de edad, los que son respondidos por el cuidador principal del niño/a.
- **Formulario F3:** en la cuarta visita se aplica la Entrevista Diagnóstica Internacional Compuesta a todas las personas de la submuestra CIDI,

este cuestionario es un instrumento que ha sido aplicado en estudios epidemiológicos en población general, y presenta validez en múltiples estudios internacionales y en Chile (Vielma *et al.*, 2002). Esta entrevista permite medir la prevalencia de trastornos mentales, y evaluar el uso de medicamentos en personas con trastornos mentales.

En esta investigación, los formularios 1 y 2 resultaron ser los más pertinentes para lograr el objetivo propuesto. El formulario 1 está conformado por 21 módulos, los cuales entregaron información de los factores conductuales y sociodemográficos. Los datos pertenecientes al factor sociodemográfico y económico se obtuvieron de los módulos I y XXI. Los datos asociados al factor conductual se obtuvieron de los módulos III, IV, XI, XII, XIII, XIV. Los datos relacionados al factor psicosocial se recogieron del módulo IX, XX, y XXI. Para obtener información sobre las enfermedades se utilizaron los módulos XVI y XVII. El formulario 2 está conformado por 14 módulos, los cuales entregaron resultados de los exámenes de laboratorio, asimismo, el módulo VII de este cuestionario considera el consumo de alcohol, el cual es parte del factor conductual.

La información que recogió cada módulo del formulario 1, y algunos del formulario 2 se encuentran en la sección de Anexos-A. A continuación, en el Cuadro 3.1.2 se encuentran las variables consideradas en el modelo y su respectiva fuente en los formularios de la ENS.

Cuadro 3.1.2: Variables utilizadas en los modelos y su respectiva fuente en los formularios de la ENS 2016-2017.

Variable	Pregunta (Formulario)
Factor demográfico y socioeconómico	
Sexo	Sexo (F1)
Estado civil	as4 (F1)
Edad	Edad (F1)
Zona geográfica	Región (F1)
Área geográfica	Área geográfica (F1)
Nivel educacional	as7_corr (F1)
Situación laboral	as10 (F1)
Sistema previsional	as5 (F1)
Factor conductual	
Consumo de alcohol	m7p9 (F2)
Consumo de cigarrillos	ta3 (F1)
Consumo agua diario	die11 (F1)
Dieta	die1a, die2, die5, die4, die6 y die8 (F1)
Actividad física	a17 (F1)
Horas de sueño diarias	ts5 (F1)
Factor psicosocial	
Estrés	ps7 (F1)
Estrés financiero	ps8 (F1)
Enfermedades	
Hipertensión	h2 (F1)
Diabetes	di3 (F1)
Exámenes de laboratorio	
Examen de sangre	m3p3 (F2)

Fuente: Elaboración propia en base a formularios de la III Encuesta Nacional de Salud 2016-2017 (EPI, 2021).

Las variables presentes en el Cuadro 3.1.2 se obtuvieron luego de un proceso previo en la base de datos entregada por el MINSAL. En primer lugar, la base original consta de 1158 variables (formulario 1 y 2) con 6233 observaciones, cada variable se debió relacionar de forma adecuada con la pregunta respectiva en los formularios. El proceso de construcción de la nueva base de datos, se compone de 4 pasos:

1. Eliminar las variables irrelevantes para esta investigación, por ejemplo, identificador de encuesta, fecha de inicio de encuesta, fecha de cierre, etc.

2. Limpieza de información, eliminar observaciones donde no se cuenta con la medición del peso o talla, ya que son los componentes del índice de masa corporal.
3. Ordenar la información, breve análisis descriptivo para observar como estaban conformadas cada una de las variables.
4. Recategorizar las variables, donde las categorías estuvieran desbalanceadas se creó una nueva variable con el objetivo de disminuir la cantidad de categorías y que éstas fueran distribuidas de forma balanceada.

Así la nueva base de datos quedó construida por 5269 observaciones y 21 variables independientes más la variable dependiente.



3.2. Técnicas Estadísticas

El análisis y modelado de los datos se llevó a cabo en tres etapas. En una primera etapa, se realizó un análisis estadístico descriptivo de los datos, pruebas de asociación bivariada entre la variable dependiente y las variables independientes del modelo, como la prueba Chi-cuadrado. Estos análisis descriptivos permiten contribuir a la comprensión de los resultados en la parte analítica.

En la segunda etapa se utilizó bosque aleatorio, en específico el criterio de importancia, para realizar una preselección de los principales determinantes de la prevalencia del sobrepeso y obesidad en Chile, así como de las posibles interacciones existentes entre las variables. El objetivo de esta etapa es facilitar la especificación del modelo de regresión logística multinomial.

En la tercera etapa, y considerando el objetivo de este estudio, se definió un modelo de regresión logística multinomial para analizar e interpretar la relación del sobrepeso y obesidad, con un conjunto de variables explicativas. Este modelo contiene variables relacionadas al factor demográfico y socioeconómico, factor conductual, factor psicosocial, enfermedades y exámenes de laboratorio.

Para la ejecución de esta investigación se utilizaron los software IBM SPSS Statistics (Versión 25.0) y Jupyter Notebooks (Versión 6.3.0). A continuación se describen las técnicas utilizadas en la presente investigación.

Machine Learning

El Machine Learning o aprendizaje automático, proviene de varias décadas atrás, y se basa fundamentalmente en la estadística. El aprendizaje automático consiste en codificar programas que ajustan automáticamente su rendimiento en función de la información de los datos entregados, esto es, mediante la realización de un modelo parametrizado, donde los parámetros se van ajustando de forma automática en función de distintos criterios de rendimiento (Igual y Seguí, 2017).

Generalmente este tipo de aprendizaje es utilizado en robótica, diagnósticos médicos, mejora de motores de búsqueda, en detección de fraude en el uso de tarjetas de crédito, asistencia virtual, aplicaciones de GPS o marketing predictivo.

Actualmente, el aprendizaje automático se puede considerar como un subcampo de la inteligencia artificial (IA) (Igual y Seguí, 2017). La IA se divide en tres clases, de acuerdo al tipo de aprendizaje:

- Aprendizaje Reforzado
- Aprendizaje Supervisado
- Aprendizaje No Supervisado

Las técnicas utilizadas en este trabajo pertenecen al aprendizaje supervisado, en específico a la clasificación del estado nutricional de personas en Chile. El aprendizaje supervisado consiste en que el algoritmo aprenda a partir de un conjunto de datos etiquetados (Igual y Seguí, 2017), es decir, se necesita intervención humana para etiquetar, clasificar e introducir los datos. Ejemplos de aprendizaje supervisado son las máquinas de soporte vectorial, árboles de decisión, bosque aleatorios, k vecinos más cercanos, entre otras. En esta investigación se

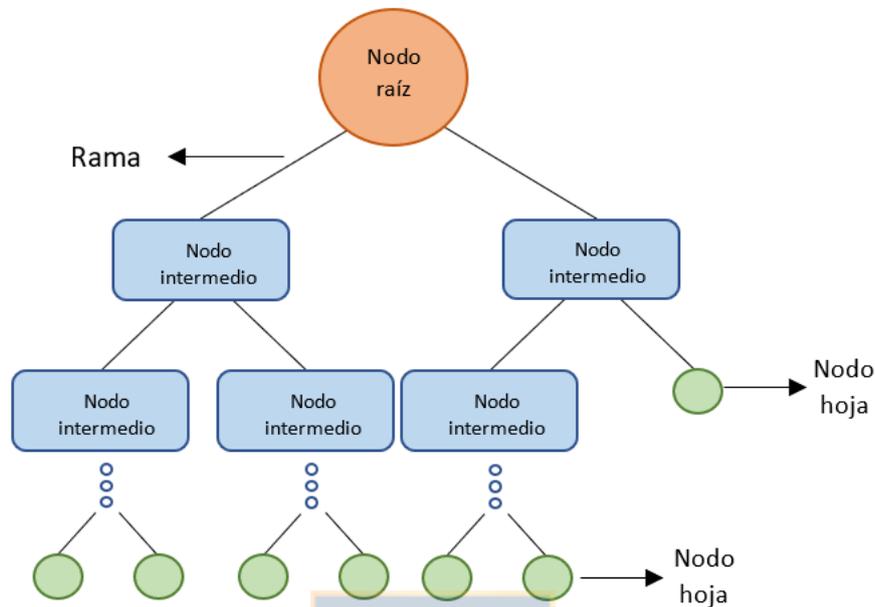
hará uso de bosque aleatorio, por lo tanto, es importante definir previamente que es un árbol de decisión para luego introducir el bosque aleatorio.

3.2.1. Árboles de Decisión

Un árbol de decisión es un método de aprendizaje considerado sencillo y eficiente para extraer conocimiento de los datos (Ertel, 2017). Un árbol de decisión se puede utilizar para regresión o clasificación, este último es nuestro caso. A diferencia de otros métodos, en el caso de árboles de decisión la información que entrega es completa, en primer lugar porque son modelos predictivos con alta precisión y estabilidad, y en segundo lugar, por la facilidad de interpretar sus resultados dada la forma de árbol (Ertel, 2017).

Un árbol de decisión se presenta como una estructura de árbol (Figura 3.2.1), y se compone principalmente de 3 elementos. En primer lugar, el nodo de decisión, indica que se debe tomar una decisión en ese momento del proceso, al nodo de decisión superior generalmente se le llama nodo raíz y a los siguientes se les denomina nodos intermedios. Luego las ramas, muestran las distintas opciones o caminos disponibles después de tomar una decisión. Y por último, los nodos hoja, terminal o de probabilidad, se encuentran en el extremo inferior de la estructura, representan los resultados, es decir, en ese punto ocurre un suceso aleatorio con una probabilidad asociada.

Para la generación de un árbol de decisión se utiliza el algoritmo de Hunt. Donde si todos los registros del conjunto de entrenamiento pertenecen a una misma clase, entonces el nodo t es un nodo hoja, en caso contrario, si los registros están distribuidos en más de una clase, se debe escoger una variable para la división de los datos en subconjuntos más reducidos. Este procedimiento se aplica de forma recursiva para cada subconjunto de los datos.

Figura 3.2.1: Diagrama árbol de decisión.

Fuente: Elaboración propia.

El algoritmo se lleva a cabo de manera que la división realizada en cada etapa sea óptima de acuerdo con algún criterio de impureza. En este caso los criterios de impureza son la entropía, índice de Gini y el error de clasificación, donde este último no fue utilizado en esta investigación.

La entropía mide que tan bien un atributo separa las observaciones de acuerdo a la clasificación deseada, es decir, la homogeneidad de la variable. Si un nodo de decisión es puro su entropía es cero y sólo tiene observaciones de una clase, mientras que si la entropía es más cercana a uno, el nodo es más impuro y se logra un balance en la frecuencia de observaciones para cada clase de la variable respuesta. La entropía se define como (Ertel, 2017):

$$H(t) = - \sum_j p(j|t) \cdot \log_2 p(j|t) \quad (3.2.1)$$

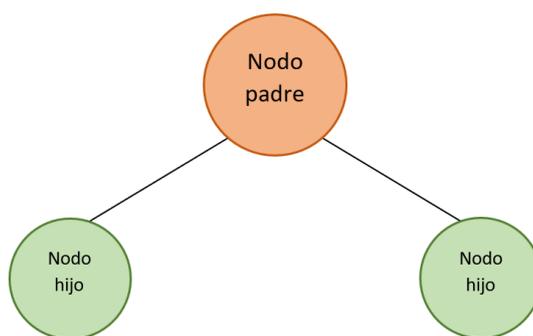
Donde t representa el nodo al cual se le calcula la entropía, y $p(j|t)$ es la probabilidad de pertenecer a la clase j de la variable respuesta estando en el nodo t .

Por otro lado, el índice de Gini mide el grado de pureza de un nodo de decisión. Toma valores entre cero y uno, donde cero representa pureza en la clasificación, es decir, todos los elementos pertenecen sólo a una clase. El valor uno indica que hay una distribución aleatoria de las observaciones en las distintas clases de la variable respuesta. El índice de Gini se define como:

$$GINI(t) = 1 - \sum_j (p(j|t))^2 \quad (3.2.2)$$

Donde los elementos t , j y $p(j|t)$ representan lo mismo que en la ecuación 3.2.1. Para la construcción del árbol de decisión, en cada nueva división se debe comparar el grado de impureza del nodo padre, respecto al nodo de impureza de los nodos hijos (Figura 3.2.2).

Figura 3.2.2: Ejemplo nodo padre y nodo hijo.



Fuente: Elaboración propia.

Esta comparación se calcula mediante la información ganada (IG), que es la resta de la impureza del nodo padre menos la media ponderada de las impurezas de los nodos hijos. Es decir:

$$IG = I(T) - \sum_{i=1}^k \frac{n_i}{n} I(i) \quad (3.2.3)$$

Donde I es el criterio de impureza utilizado, que puede ser entropía, índice de Gini o el error de clasificación, T es el nodo padre, i representa a cada nodo hijo, n_i es la cantidad de observaciones en el nodo i , y n es el total de observaciones, es decir, $\sum_{i=1}^k n_i$. Se selecciona la variable que mayor ganancia en la información entrega, esto es, cuando la media ponderada de las impurezas de los nodos hijos es mínima.

3.2.2. Bosque Aleatorio

Bosque Aleatorio o Random Forest (RF) es un método de consenso, ya que se basa en la combinación de distintos árboles de decisión, utilizando la técnica de agregación llamada voto por mayorías (Igual y Seguí, 2017). Para este método se utilizan dos parámetros, el número de árboles que generalmente se ajusta mediante validación cruzada, y el número de atributos que cada árbol utilizará.

El procedimiento de este método se puede resumir en (Igual y Seguí, 2017):

- Dividir los datos iniciales en varios subconjuntos compuestos de muestras aleatorias, el muestreo se realiza con la técnica de bootstrapping.
- Aplicar y entrenar un modelo de árbol de decisión para cada subconjunto obtenido del punto anterior.
- Establecer un criterio de consenso de todos los resultados obtenidos, generalmente se utiliza el consenso de quien obtenga la mayor cantidad de votos.

Este método se considera fácil y rápido de usar, debido a que se necesita poco preprocesamiento de los datos, estos no deben ser necesariamente reescalados o

transformados. Es un buen método cuando se tiene una gran cantidad de datos y variables, y se entrena de manera rápida.

Los clasificadores de RF al igual que el árbol de decisión, permite revelar la importancia de las variables utilizadas, a través de la medida de impureza (entropía, índice de Gini o error de clasificación). En cada árbol de decisión, las variables reciben una puntuación de acuerdo con la medida de impureza del nodo de decisión. La importancia de una variable dentro un bosque aleatorio se mide como la media de las puntuaciones obtenidas en cada árbol, que luego se normaliza por 1 (Ertel, 2017).

En este caso la importancia de las variables dentro del bosque aleatorio son un aporte en la construcción del modelo de regresión logística multinomial, dado que nos dará el orden en que se deben ingresar las variables al modelo.

La variable más importante dentro del bosque aleatorio será la que en primer lugar se ingrese con todas sus categorías (a excepción de la categoría de referencia) al modelo de regresión. Se evalúa la significancia de cada una de las categorías y se seleccionan las que mejor aportan al modelo, luego se ingresa la segunda variable más importante de acuerdo a los resultados del bosque aleatorio, se observa el cambio en los indicadores de ajuste y el incremento del poder explicativo, si estos indicadores no son favorables, la variable se excluye y se continúa el proceso con la siguiente variable más importante. Así se repite el proceso hasta alcanzar el conjunto de variables óptimo que logre maximizar el poder explicativo del modelo de regresión logística multinomial.

Regresión Logística

El análisis de regresión logística es una técnica estadística multivariable destinada al análisis de una relación de dependencia entre una variable dependiente categórica y un conjunto de variables independientes, de forma similar al análisis de regresión lineal. El objetivo, en este caso, es poder estimar las probabilidades de un suceso definido por la variable dependiente cualitativa dicotómica o con más categorías, en función de un conjunto de variables predictoras cualitativas o cuantitativas (López-Roldán y Fachelli, 2016).

Por sus características, los modelos de regresión logística permiten dos finalidades: i) Cuantificar la importancia de la relación existente entre cada una de las covariables y la variable dependiente, lo que lleva implícito también clarificar la existencia de interacción entre covariables respecto a la variable dependiente, es decir, conocer la odds ratio para cada covariable; ii) Clasificar individuos dentro de las categorías (presente/ausente) de la variable dependiente, según la probabilidad que tenga de pertenecer a una de ellas, dada la presencia de determinadas covariables.

Para López-Roldán y Fachelli (2016), la regresión logística tiene dos modalidades:

- Regresión logística binaria: cuando se desea explicar un suceso dicotómico (presencia/ausencia).
- Regresión logística multinomial: cuando se quiere explicar una variable cualitativa politómica. Para esto se requiere convertir la variable en diversas variables dicotómicas ficticias, es decir, crear tantas variables dicotómicas (dummy) como categorías tenga la variable menos una, la que hará de categoría de referencia.

Para la regresión logística multinomial, se diferencia la situación en que la variable a explicar es politómica nominal, o politómica ordinal.

3.2.3. Regresión Logística Multinomial

La regresión logística multinomial es una extensión multivariante de la regresión logística binaria, y es utilizada en modelos con variable dependiente de tipo nominal con más de dos categorías, las variables independientes pueden ser continuas o categóricas. En este caso la variable dependiente se definió en tres categorías: i) persona con peso normal o delgado, ii) persona con sobrepeso y, iii) persona con obesidad.

En primer lugar se debe verificar los supuestos del modelo:

1. Variables predictoras deben ser categóricas o continuas, y no requieren supuesto de normalidad.
2. No se requiere linealidad, es decir, la variable dependiente debe ser categórica con dos o más niveles.
3. Independencia del error, debe ser un estudio intergrupacional.
4. No debe existir multicolinealidad entre las variables.

En un modelo de regresión logística multinomial, donde se considera la variable dependiente Y categórica nominal politómica con soporte $\{1, 2, 3\}$, representa el estado nutricional con 1: normal o bajo peso; 2: sobrepeso y 3: obeso. Y con X_1, X_2, \dots, X_n variables independientes, la ecuación es la siguiente:

$$P(Y = i | X_1, \dots, X_n) = \frac{\exp[\beta_{0i} + \beta_{1i}X_1 + \dots + \beta_{ni}X_n]}{1 + \sum_{j=1}^3 \exp[\beta_{0j} + \beta_{1j}X_1 + \dots + \beta_{nj}X_n]} \quad (3.2.4)$$

Donde

- i : categoría del estado nutricional que se quiere calcular la probabilidad.
- β : parámetros que se desean estimar.
- j : 1,2 y 3; representa los diferentes niveles del estado nutricional.

La ecuación 3.2.4, se interpreta como la probabilidad de que un individuo pertenezca a la categoría i , dado un conjunto de covariables. Este método permite modelar una relación entre las variables explicativas y la variable dependiente, y los resultados obtenidos permiten analizar el impacto de las distintas variables explicativas, tanto por su magnitud como por el signo. De esta forma, es posible establecer qué variables tienen un mayor impacto en el sobrepeso y la obesidad, y si ese impacto es positivo o negativo.

Para la determinación del mejor modelo explicativo, se realizaron los siguientes pasos:

1. Selección de las variables del modelo: esto se efectuó en primera instancia, mediante un análisis exploratorio de los datos para buscar relaciones existentes entre la variable respuesta y las variables explicativas, y se complementó con la prueba estadística Chi-cuadrado. En este paso también se realizaron recategorizaciones a las variables que presentaron un alto número de categorías.
2. Estimación de los coeficientes: los coeficientes determinan la importancia que cada variable independiente tiene en la explicación de la variable dependiente, para este proceso de estimación se aplica el método de máxima verosimilitud.
3. Evaluación del modelo: se refiere a establecer la capacidad explicativa del modelo o bondad de ajuste, para lo cual se utilizó el estadísticos R^2 de McFadden. Por otro lado, la capacidad discriminatoria del modelo se estableció mediante la matriz de clasificación, donde se observa el porcentaje de casos bien clasificados.

Una característica interesante de la regresión logística es la relación que guarda con un parámetro de cuantificación de riesgo, conocido como *odds ratio* (OR). El odds asociado a un suceso se define como el cociente entre la probabilidad de que ocurra el suceso con relación a la probabilidad de que no ocurra, es decir, $odds = \frac{p}{1-p}$, con p la probabilidad de ocurrencia del suceso.

El odds ratio se define como un cociente de odds, es una medida de asociación entre la exposición y un resultado. El OR representa la probabilidad de que un resultado ocurra dada una exposición particular, comparada con la probabilidad de que el resultado ocurra en ausencia de esa exposición. Un $OR > 1$ indica un aumento en la ocurrencia del evento dada la exposición, mientras que un $OR < 1$ indica una disminución en la ocurrencia del evento, también frente a la exposición.



Capítulo 4

Resultados

4.1. Análisis descriptivo

En primer lugar se realiza un análisis descriptivo para la caracterización de la muestra, (Cuadro 4.1.1). La base de datos de la ENS 2016-2017 tiene 6233 observaciones, de las cuales solo 5269 contienen información sobre mediciones biofisiológicas y, por lo tanto, del índice de masa corporal, medición clave para clasificar a los individuos por su estado nutricional.

Cuadro 4.1.1: Caracterización de la muestra basada en el estado nutricional.

Variable (n)	Categoría	Porcentaje muestra	Porcentaje dentro del estado nutricional		
			Normal	Sobrepeso	Obeso
Sexo (5269)***	Mujer	63,9 %	61,9 %	59,1 %	70,0 %
	Hombre	36,1 %	38,1 %	40,9 %	30,0 %
Estado civil (5269)***	Soltero	51,0 %	61,8 %	49,3 %	46,3 %
	Casado o en pareja	49,0 %	38,2 %	50,7 %	53,7 %
Edad (5269)***	18-34 años	24,0 %	35,8 %	22,4 %	18,7 %
	35-50 años	24,7 %	20,4 %	24,9 %	27,0 %
	51-64 años	25,8 %	17,8 %	27,6 %	28,8 %
	65 o más años	25,4 %	40,4 %	32,3 %	31,3 %
Zona geográfica (5269)***	Zona norte	26,0 %	28,4 %	27,1 %	23,5 %
	Zona centro	22,1 %	24,4 %	22,0 %	20,8 %
	Región Metropolitana	15,0 %	15,0 %	16,7 %	13,3 %
	Zona sur	26,4 %	21,8 %	24,8 %	30,7 %
	Zona austral	10,6 %	10,4 %	9,5 %	11,7 %

(Continúa en la página siguiente)

Variable (n)	Categoría	Porcentaje muestra	Porcentaje dentro del estado nutricional			
			Normal	Sobrepeso	Obeso	
Área geográfica (5269)***	Urbana	83,9 %	86,9 %	84,9 %	81,1 %	
	Rural	16,1 %	13,1 %	15,1 %	18,9 %	
Nivel educacional (5223)***	Hasta educación básica	32,5 %	25,0 %	29,3 %	40,3 %	
	Hasta educación media o técnica	34,9 %	35,3 %	36,1 %	33,6 %	
	TNS-Profesional-Postgrado	32,5 %	39,7 %	34,7 %	26,1 %	
Situación laboral (5251)***	Trabajando	47,8 %	44,1 %	53,0 %	44,7 %	
	Jubilado	20,5 %	20,5 %	19,9 %	21,2 %	
	Otros	31,7 %	35,4 %	27,1 %	34,0 %	
Sistema previsional (5213)***	FONASA	84,6 %	81,8 %	83,6 %	87,3 %	
	FFAA-ISAPRE-Otra-Ninguno	15,4 %	18,2 %	16,4 %	12,7 %	
Consumo alcohol (5269)***	Nunca	35,2 %	34,5 %	33,0 %	37,9 %	
	1 vez al mes	36,4 %	31,3 %	37,2 %	38,7 %	
	2 a 4 veces al mes	20,6 %	24,0 %	21,3 %	17,9 %	
Consumo cigarrillos (5269)***	2 o más veces a la semana	7,7 %	10,2 %	8,5 %	5,6 %	
	No fuma	46,8 %	48,7 %	45,1 %	47,4 %	
	Ha dejado de fumar	24,4 %	19,7 %	24,8 %	26,8 %	
Consumo agua diario (5264)***	Si fuma	28,8 %	31,6 %	30,1 %	25,8 %	
	Ningun vaso	7,9 %	8,2 %	8,9 %	6,6 %	
	1-4 vasos	57,2 %	59,7 %	57,3 %	55,6 %	
Dieta (5269)	5-8 vasos	28,0 %	26,3 %	27,0 %	30,1 %	
	Más de 8 vasos	6,9 %	5,8 %	6,8 %	7,7 %	
	Mala	45,7 %	45,9 %	45,4 %	46,0 %	
Actividad física (5269)***	Buena	54,3 %	54,1 %	54,6 %	54,0 %	
	Sí	10,2 %	14,3 %	10,5 %	7,3 %	
	A veces	10,6 %	13,3 %	11,5 %	8,2 %	
Horas de sueño diarias (5269)**	No	79,2 %	72,4 %	78,0 %	84,5 %	
	1-5 horas	13,5 %	11,5 %	13,0 %	15,2 %	
	6-8 horas	73,1 %	73,7 %	73,7 %	72,1 %	
Estrés (5268)	Más de 8 horas	13,4 %	14,8 %	13,3 %	12,7 %	
	Varias veces	21,4 %	22,0 %	21,8 %	20,6 %	
	Algunas veces	39,3 %	38,5 %	40,7 %	38,3 %	
Estrés financiero (5212)***	Nunca	39,4 %	39,6 %	37,5 %	41,1 %	
	Moderado o alto	54,2 %	49,2 %	55,7 %	55,6 %	
	Poco o nada	45,8 %	50,8 %	44,3 %	44,4 %	
Hipertensión (5269)***	Sí	33,9 %	22,4 %	31,3 %	43,2 %	
	No	66,1 %	77,6 %	68,7 %	56,8 %	
Diabetes (5268)***	Sí	15,1 %	8,6 %	12,9 %	21,3 %	
	No	84,9 %	91,4 %	87,1 %	78,7 %	
Colesterol LDL (3511)***	Valor mayor que normal	53,5 %	43,5 %	56,6 %	56,4 %	
	Valor normal (<100mg/dL)	46,5 %	56,5 %	43,4 %	43,6 %	
Triglicéridos (3536)***	Valor mayor que normal	35,5 %	18,3 %	35,8 %	45,7 %	
	Valor normal (<150mg/dL)	64,5 %	81,7 %	64,2 %	54,3 %	
Glucosa (4887)***	Valor mayor que normal	26,6 %	13,2 %	22,1 %	39,1 %	
	Valor normal (<100mg/dL)	73,4 %	86,8 %	77,9 %	60,9 %	
Total (n)			1206	2034	2029	
Notas: Test Chi-cuadrado * $p < 0,1$; ** $p < 0,05$; *** $p < 0,01$.			Porcentaje	22,9 %	38,6 %	38,5 %

Fuente: Elaboración propia.

Comenzando con el **factor demográfico y socioeconómico**, el 63,9% de los encuestados son mujeres, y el 36,1% son hombres. Se obtuvo mayor porcentaje de sobrepeso y obesidad mujeres. En cuanto al estado civil, 2582 personas declaran estar casados o en pareja, y estos presentaron mayor porcentaje dentro de las personas con sobrepeso u obesidad, con 50,7% y 53,7% respectivamente.

La edad se categorizó en 4 grupos, siendo el grupo de edad de 65 años o más el que tiene un mayor porcentaje dentro del estado nutricional sobrepeso y obeso. Se creó la variable zona geográfica a partir de la región de residencia declarada, dada que la región metropolitana estaba ampliamente representada es que se consideró como una zona geográfica, la zona norte y sur son las que contienen mayor número de personas encuestadas con 1371 y 1389 respectivamente. Con respecto al área geográfica de los encuestados, un 83,9% pertenecen a la zona urbana, y un 16,1% a la zona rural.

En cuanto al nivel educacional, las personas con nivel educacional básico completo o menor, presentaron mayor porcentaje dentro del estado nutricional obeso, con 40,3%. Para el nivel de educación media o técnico completo, el mayor porcentaje se presentó en sobrepeso con 36,1%, y para quienes presenta un nivel educacional de técnico nivel superior, profesional o postgrado completo, el mayor porcentaje se encontró en las personas con peso normal o delgados con 39,7%. Con respecto a la situación laboral de los encuestados, el 47,8% se encuentra trabajando, el 20,5% está jubilado y el 31,7% restante se encuentra estudiando, buscando trabajo o en otras situaciones. En cuanto al sistema previsional, cerca del 85% de los encuestados pertenecen al Fondo Nacional de Salud (FONASA), los restantes se encuentran en ISAPRE, FFAA, u otras.

Con respecto al **factor conductual**, cerca de un tercio de la muestra declaró

no consumir alcohol (35,2%), mientras que los dos tercios restantes declararon consumir por lo menos una vez al mes. En cuanto al consumo de cigarrillos, cerca del 50% declaró no fumar, dentro de las personas con sobrepeso el 45,1% no fuma, y el 30,1% si lo hace, en las personas con obesidad las cifras son similares, el 47,4% no fuma, y un 25,8% fuma. El consumo de agua diario se cuantificó en vasos de 200 mililitros, el 57,2% declaró consumir entre 1 a 4 vasos diarios, y sólo un 6,9% declaró consumir más de 8 vasos de agua diarios. Con respecto a la dieta se calculó de acuerdo al consumo de frutas, verduras, pescados, cereales y legumbres, poco más del 50% de los encuestados tienen una buena dieta. En cuanto realización de actividad física con duración de a lo menos 30 minutos en la semana, sólo el 10,2% lo hace frecuentemente, 10,6% declararon realizar actividad física a veces, mientras que el 79,2% no realiza actividad física. Por último con respecto a los hábitos de sueño el 73,1% de los encuestados duerme entre 6 a 8 horas diarias durante los días de semana.

Por otro lado, para el **factor psicosocial** se consideraron las variables estrés y estrés financiero. De los 5268 encuestados que respondieron la pregunta relacionada al estrés sólo el 21,4% declaró sentirse estresado varias veces, el 39,3% se sintió estresado algunas veces, representan el 40,7% de las personas con sobrepeso. El estrés financiero que se produce a consecuencia de distintos problemas económicos, se hizo presente en el 54,2% de los encuestados, y representan al 55,7% y 55,6% de las personas con sobrepeso y obesidad respectivamente.

Con respecto a las **enfermedades** que han sido diagnosticadas por el personal de salud, el 33,9% y 15,1% de los encuestados fueron diagnosticados con hipertensión y diabetes, respectivamente.

En cuanto a los **exámenes de laboratorios**, se consideraron colesterol LDL, triglicéridos y glucosa. De los exámenes mencionados, triglicéridos y glucosa presentaron un porcentaje mayor a 50% en la categoría valor normal, a excepción del colesterol LDL con 46,5% en la categoría valor normal (<100mg/dL).

El resumen gráfico de las variables mencionadas se encuentra en Anexo-B.



4.2. Selección de variables

Para realizar la selección de variables mediante el criterio de importancia se utilizó bosques aleatorios. Se entrenaron 14 modelos de bosques aleatorios con 100, 200, 400, 800, 1000 y 1200 árboles, utilizando los criterios de selección de gini y entropía, y profundidad máxima de 20. Se estimó la precisión de cada modelo mediante validación cruzada con 5 grupos. El tamaño de muestra es $n = 3301$, dado que algunas variables presentaron observaciones vacías. Los resultados del entrenamiento se encuentran en el Cuadro 4.2.1.

Cuadro 4.2.1: Resultados entrenamiento bosques aleatorios.

Criterio selección	Número de árboles	Precisión	SD
Entropía	600	44,11 %	0,0217
Entropía	1000	43,83 %	0,0235
Gini	1000	43,74 %	0,0214
Gini	1200	43,65 %	0,0240
Entropía	1200	43,56 %	0,0228
Gini	400	43,44 %	0,0190
Gini	600	43,41 %	0,0178
Gini	100	43,29 %	0,0247
Entropía	800	43,26 %	0,0165
Entropía	100	43,14 %	0,0232
Entropía	200	43,14 %	0,0178
Entropía	400	43,11 %	0,0226
Gini	800	42,96 %	0,0180
Gini	200	42,53 %	0,0210

Fuente: Elaboración propia.

Se seleccionó el modelo de bosque aleatorio utilizando criterio de selección de entropía con 600 árboles de decisión para la posterior realización del análisis del criterio de importancia de las variables. El ranking de las variables de acuerdo al índice de importancia se encuentra en el Cuadro 4.2.2.

Cuadro 4.2.2: Importancia de variables a través de bosque aleatorio.

Ranking	Variable	Índice de importancia
1	Zona geográfica	0,0953
2	Consumo alcohol	0,0764
3	Sexo x Estado civil	0,0758
4	Estrés	0,0747
5	Edad	0,0684
6	Consumo agua diario	0,0676
7	Consumo cigarrillos	0,0635
8	Nivel educacional	0,0572
9	Situación laboral	0,0538
10	Horas de sueño diarias	0,0503
11	Dieta	0,0440
12	Estrés financiero	0,0404
13	Colesterol LDL	0,0362
14	Actividad física	0,0349
15	Triglicéridos	0,0317
16	Glucosa	0,0309
17	Área geográfica	0,0274
18	Hipertensión	0,0273
19	Sistema previsional	0,0250
20	Diabetes	0,0192

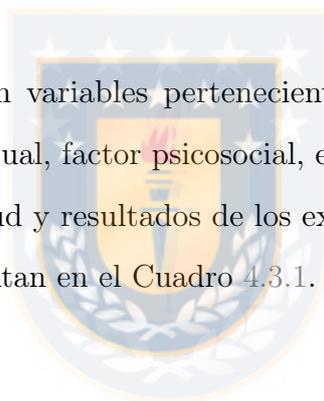
Fuente: Elaboración propia.

Los resultados que entrega esta herramienta son de ayuda para la etapa de estimación de un modelo de regresión logística multinomial, ya que da una noción del orden en que son añadidas al modelo.

4.3. Regresión Logística Multinomial

Para modelar el sobrepeso y la obesidad, a través un modelo de regresión logística multinomial, debemos tener claridad de la variable dependiente, como se mencionó anteriormente, está compuesta por tres categorías; i) persona con peso normal o delgado, ii) persona con sobrepeso y, iii) persona con obesidad. En este caso la categoría “persona con peso normal o delgado” se estableció como categoría base o de referencia. Para la estimación del modelo logístico multinomial, el orden de elección de variables a introducir se realizó a partir de los resultados de bosques aleatorios.

El modelo se estimó con variables pertenecientes al factor demográfico y socioeconómico, factor conductual, factor psicosocial, enfermedades diagnosticadas por un profesional de la salud y resultados de los exámenes de laboratorio. Las variables utilizadas se presentan en el Cuadro 4.3.1.



Cuadro 4.3.1: Variables utilizadas en estimación del modelo logístico multinomial para el estado nutricional.

Variable dependiente	Categorías
<i>Estado Nutricional</i>	1. Sobrepeso 2. Obeso 3. Normal ^{RC}
Variables independientes	Categorías
<i>Sexo x Estado civil</i>	1. Mujer casada o en pareja 2. Mujer soltera 3. Hombre casado o en pareja 4. Hombre soltero ^{RC}
<i>Edad</i>	1. 18-34 años 2. 35-50 años 3. 51-64 años 4. 65 o más años ^{RC}
<i>Zona geográfica</i>	1. Zona norte 2. Zona centro 3. Región Metropolitana 4. Zona sur 5. Zona austral ^{RC}
<i>Área geográfica</i>	1. Urbana 2. Rural ^{RC}
<i>Nivel educacional</i>	1. Hasta educación básica 2. Hasta educación media o técnica 3. TNS-Profesional-Postgrado ^{RC}
<i>Situación laboral</i>	1. Trabajando 2. Jubilado 3. Otros ^{RC}
<i>Sistema previsional</i>	1. FONASA 2. FFAA-ISAPRE-Otra-Ninguna ^{RC}
<i>Consumo alcohol</i>	1. Nunca 2. 1 vez al mes 3. 2 a 4 veces al mes 4. 2 o más veces a la semana ^{RC}
<i>Consumo cigarrillos</i>	1. No fuma 2. Ha dejado de fumar 3. Si fuma ^{RC}
<i>Consumo agua diario</i>	1. Ningún vaso 2. 1-4 vasos 3. 5-8 vasos 4. Más de 8 vasos ^{RC}

(Continúa en la página siguiente)

Variables independientes	Categorías
<i>Dieta</i>	1. Mala 2. Buena ^{RC}
<i>Actividad física</i>	1. Sí 2. A veces 3. No ^{RC}
<i>Horas de sueño diarias</i>	1. 1-5 horas 2. 6-8 horas 3. Más de 8 horas ^{RC}
<i>Estrés</i>	1. Varias veces 2. Algunas veces 3. Nunca ^{RC}
<i>Estrés financiero</i>	1. Moderado o alto 2. Poco o nada ^{RC}
<i>Hipertensión</i>	1. Sí 2. No ^{RC}
<i>Diabetes</i>	1. Sí 2. No ^{RC}
<i>Colesterol LDL</i>	1. Valor mayor que normal 2. Valor normal (<100mg/dL) ^{RC}
<i>Triglicéridos</i>	1. Valor mayor que normal 2. Valor normal (<150mg/dL) ^{RC}
<i>Glucosa</i>	1. Valor mayor que normal 2. Valor normal (<100mg/dL) ^{RC}

Fuente: Elaboración propia.

Los resultados de la estimación del modelo se muestran en el Cuadro 4.3.2. El modelo es globalmente significativo, con McFadden pseudo $R^2 = 0,091$. Con respecto a los resultados para cada estado nutricional, se deben interpretar siempre en relación con la categoría de referencia "persona con peso normal o delgado".

Cuadro 4.3.2: Resultados estimación del modelo logístico multinomial para el estado nutricional.

Variables	Sobrepeso	Obeso
	β (Desv. Error)	β (Desv. Error)
Constante	-0,9032** (0,4234)	-2,4895*** (0,4568)
Factor demográfico y socioeconómico		
<i>Sexo x Estado civil</i>		
Mujer casada o en pareja	0,4113*** (0,1555)	1,0114*** (0,1684)
Mujer soltera	0,0759 (0,1415)	0,4858*** (0,1567)
Hombre casado o en pareja	0,5409*** (0,1575)	0,5960*** (0,1750)
<i>Edad</i>		
18-34 años	0,2840 (0,2054)	1,1252*** (0,2219)
35-50 años	0,5341*** (0,198)	1,3245*** (0,2103)
51-64 años	0,3916** (0,1793)	0,6158*** (0,1893)
<i>Zona geográfica</i>		
Zona norte	-0,0161 (0,1796)	-0,4885*** (0,1865)
Zona centro	-0,1704 (0,1829)	-0,5402*** (0,1901)
Región Metropolitana	0,2320 (0,1960)	-0,2594 (0,2067)
Zona sur	0,0850 (0,1846)	-0,0967 (0,1894)
<i>Área geográfica</i>		
Urbana	-0,1891 (0,1452)	-0,0939 (0,1508)
<i>Nivel educacional</i>		
Hasta educación básica	0,1354 (0,1474)	0,5247*** (0,1550)
Hasta educación media o técnica	0,0543 (0,1186)	0,147 (0,1279)
<i>Situación laboral</i>		
Trabajando	0,1810 (0,1185)	0,0178 (0,1253)
Jubilado	0,2323 (0,1857)	0,4742** (0,1936)
<i>Sistema previsional</i>		
FONASA	0,1250 (0,1346)	0,2722* (0,1498)
Factor conductual		
<i>Consumo alcohol</i>		
Nunca	0,1307 (0,1855)	0,2689 (0,2094)
1 vez al mes	0,3146* (0,1796)	0,5992*** (0,2032)
2 a 4 veces al mes	0,0818 (0,1859)	0,3741* (0,2111)
<i>Consumo cigarrillos</i>		
No fuma	0,1958* (0,1169)	0,2982** (0,1255)
Ha dejado de fumar	0,3244** (0,1370)	0,5144*** (0,1461)
<i>Consumo agua diario</i>		
Ningún vaso	-0,3255 (0,2509)	-0,9737*** (0,2729)
1-4 vasos	-0,2493 (0,1981)	-0,4903** (0,2076)
5-8 vasos	-0,1870 (0,2077)	-0,1903 (0,2167)
<i>Dieta</i>		
Mala	-0,0326 (0,0993)	0,0605 (0,1053)

(Continúa en la página siguiente)

Variables	Sobrepeso β (Desv. Error)	Obeso β (Desv. Error)
<i>Actividad física</i>		
Sí	-0,2459 (0,1537)	-0,4657*** (0,1725)
A veces	-0,1795 (0,1576)	-0,2519 (0,1725)
<i>Horas de sueño diarias</i>		
1-5 horas	0,0611 (0,1916)	0,4643** (0,2024)
6-8 horas	-0,0094 (0,1438)	0,1904 (0,1542)
Factor psicosocial		
<i>Estrés</i>		
Varias veces	0,0611 (0,1387)	-0,2750* (0,1473)
Algunas veces	0,0689 (0,1126)	-0,2018* (0,1190)
<i>Estrés financiero</i>		
Moderado o alto	0,2230** (0,1006)	0,2299** (0,1070)
Enfermedades		
<i>Hipertensión</i>		
Sí	0,3317*** (0,1201)	0,7578*** (0,1244)
<i>Diabetes</i>		
Sí	0,3349* (0,1772)	0,4423** (0,1775)
Exámenes de laboratorio		
<i>Colesterol LDL</i>		
Valor mayor que normal	0,3669*** (0,0992)	0,3682*** (0,1053)
<i>Triglicéridos</i>		
Valor mayor que normal	0,7791*** (0,1163)	1,1288*** (0,1202)
<i>Glucosa</i>		
Valor mayor que normal	0,3052** (0,1420)	1,1186*** (0,1417)

Notas: Categoría de referencia: "Peso normal o bajo"; Prueba de razón de verosimilitud: $\chi^2 = 645,937$; significancia = 0,000; McFadden pseudo $R^2 = 0,091$. * $p < 0,1$; ** $p < 0,05$; *** $p < 0,01$. Tamaño de muestra = 3301.

Fuente: Elaboración propia.

De acuerdo a los resultados de la estimación del modelo, al analizar el estado nutricional en la categoría “sobrepeso” se observa que para las variables del **factor demográfico y socioeconómico** se obtuvieron resultados positivos. La variable sexo x estado civil fue positiva para el caso de los hombres casados o en pareja, estos aumentan la probabilidad en 1,72 veces ($\exp(0,5409)$) de presentar sobrepeso, en comparación con la categoría peso normal o delgado, en el caso de las mujeres casadas o en pareja la probabilidad aumenta en 1,51 veces ($\exp(0,4113)$). La edad presentó resultados positivos en dos categorías, cuando se tiene entre 35 a 50 años aumenta la probabilidad en 1,71 veces ($\exp(0,5341)$),

mientras que entre 51 a 64 años la probabilidad de presentar sobrepeso aumenta en 1,48 veces ($\exp(0,3916)$), en comparación con la categoría de referencia.

Con respecto al **factor conductual**, sólo el consumo de alcohol y cigarrillos resultó significativo para las personas con sobrepeso. Quienes declararon consumir alcohol 1 vez al mes aumentan la probabilidad de presentar sobrepeso en 1,37 veces ($\exp(0,3146)$), en comparación con quienes presentan peso normal o bajo. En cuanto al consumo de cigarrillos quienes declararon no fumar presentan un aumento de 1,22 veces ($\exp(0,1958)$) en la probabilidad, mientras que quienes han dejado de fumar presentan un aumento mayor en la probabilidad, este es de 1,38 veces ($\exp(0,3244)$).

En el **factor psicosocial**, resultó significativo el estrés financiero en la categoría moderado o alto, es decir, aumenta en 1,25 veces ($\exp(0,2230)$) la probabilidad de presentar sobrepeso cuando la persona se siente altamente estresada por su situación financiera.

Dentro de la dimensión de **enfermedades** diagnosticada por un médico, u otro profesional de la salud, las personas diagnosticadas con hipertensión, aumentan en 1,39 veces ($\exp(0,3317)$) la probabilidad de presentar sobrepeso con respecto a la categoría de referencia, asimismo las personas diagnosticadas con diabetes presentan un aumento de 1,40 veces ($\exp(0,3349)$) en la probabilidad de presentar sobrepeso.

Por último, dentro de los **exámenes de laboratorio** resultaron todas las variables positivas y significativas. Si una persona presenta un colesterol LDL mayor o igual 100mg/dL su probabilidad de padecer sobrepeso aumenta en 1,44 veces ($\exp(0,3669)$) con respecto a quienes tienen peso normal o bajo. En

cuanto a los triglicéridos, un resultado mayor o igual a 150mg/dL aumentan la probabilidad en 2,18 veces ($\exp(0,7791)$), del mismo modo para los exámenes de glucosa, cuando se presenta un resultado mayor o igual a 100mg/dL la probabilidad de padecer sobrepeso aumenta en 1,36 veces ($\exp(0,3052)$).

En cuanto a los resultados de la estimación respecto al estado nutricional en la categoría obesidad, se observó que en el **factor demográfico y socioeconómico** resultaron significativas todas las variables, con excepción del área geográfica. Las mujeres y hombres casados o en pareja, aumentan la probabilidad de presentar obesidad en 2,75 veces ($\exp(1,0114)$) y 1,81 veces ($\exp(0,5960)$), respectivamente, en comparación con la categoría base. En cuanto a la mujer soltera, presenta un aumento en la probabilidad padecer obesidad de 1,63 veces ($\exp(0,4858)$), cifra menor en comparación a las mujeres casadas o en pareja.

Continuando con la variable edad, esta resultó ser significativa y positiva en las 3 categorías, siendo mayor el aumento de la probabilidad de padecer obesidad en las personas entre 35 y 50 años de edad, con un aumento de 3,76 veces ($\exp(1,3245)$). Para quienes tienen entre 18 y 34 años el aumento es de 3,08 veces ($\exp(1,1252)$), y para personas entre 51 a 64 años se presentó un aumento de 1,85 veces ($\exp(0,6158)$) en la probabilidad de tener obesidad con respecto a personas con peso normal o bajo. En cuanto a la zona geográfica, se observó un aumento en la probabilidad de pertenecer al estado nutricional obeso, para las personas que residen en la zona norte y centro del país, este aumento fue de 0,61 veces ($\exp(-0,4885)$) y 0,58 veces ($\exp(-0,5402)$), en cada caso. Con respecto al nivel educacional, quienes tienen educación básica completa o un menor nivel educacional, se presentó un aumento de 1,69 veces ($\exp(0,5247)$), en la probabilidad de ser obeso. En cuanto a la situación laboral de los entrevistados, quienes declararon estar jubilados tienen un aumento de

1,61 veces ($\exp(0,4742)$) en la probabilidad, asimismo quienes tienen FONASA como sistema previsional presentaron un aumento de 1,31 veces ($\exp(0,2722)$) en la probabilidad de padecer obesidad, en comparación con la categoría de referencia.

Dentro del **factor conductual**, las personas que declararon consumir alcohol 1 vez al mes presentan un aumento en la probabilidad de tener obesidad de 1,82 veces ($\exp(0,5992)$), con respecto al consumo de cigarrillos, las personas que no fuman tienen un aumento de 1,35 veces ($\exp(0,2982)$) en la probabilidad, este aumento se incrementa para quienes han dejado de fumar, teniendo un aumento de 1,67 veces ($\exp(0,5144)$) en la probabilidad de padecer exceso de peso, con respecto a quienes tienen peso normal o bajo.

Continuando con el factor conductual, el consumo de agua presentó resultados significativos y negativos, las personas que no consumen agua diariamente (ningún vaso) resultaron tener un aumento de 0,38 veces ($\exp(-0,9737)$) en la probabilidad, mientras que para quienes consumen entre 1 a 4 vasos diarios de agua el aumento es de 0,61 veces ($\exp(-0,4903)$). Quienes declararon realizar actividad física fuera del horario laboral frecuentemente, presentan un aumento de 0,63 veces ($\exp(-0,4657)$), en la probabilidad de tener obesidad en comparación con la categoría de referencia. Por último en personas que duermen entre 1 a 5 horas diarias, la probabilidad de tener obesidad aumenta 1,59 veces ($\exp(0,4643)$).

En la dimensión del **factor psicosocial**, resultó significativo el estrés en todas sus categorías, es decir, se presenció un aumento en la probabilidad de padecer obesidad de 0,76 veces ($\exp(-0,2750)$) y 0,82 veces ($\exp(-0,2018)$) en personas que declaran sentir estrés varias y algunas veces respectivamente, durante el último año. Por otro lado, aumenta en 1,26 veces ($\exp(0,2299)$) la probabilidad de presentar sobrepeso cuando la persona se siente altamente estresada por su

situación financiera.

Con respecto a las **enfermedades** diagnosticadas, se presentó un aumento de 2,13 veces ($\exp(0,7578)$) y 1,56 veces ($\exp(0,4423)$), en la probabilidad de tener obesidad, para quienes son diagnosticados con hipertensión y diabetes, respectivamente, en comparación a quienes presentaron peso normal o bajo.

Por último, en cuanto a los resultados de **exámenes de laboratorio** todas las variables resultaron ser significativas y positivas. Para quienes presentaron un colesterol LDL mayor o igual 100mg/dL la probabilidad de padecer obesidad aumenta en 1,45 veces ($\exp(0,3682)$) con respecto a quienes tienen peso normal o bajo. En cuanto a los triglicéridos, un resultado mayor o igual a 150mg/dL tiene un aumento en la probabilidad de 3,09 veces ($\exp(1,1288)$), para los exámenes de glucosa cuando se presenta un resultado mayor o igual a 100mg/dL la probabilidad de padecer obesidad aumenta en 3,06 veces ($\exp(1,1186)$).

La tabla de clasificación (Cuadro 4.3.3) indica que la precisión del modelo fue de 50,8%. Resultados importantes son la sensibilidad y especificidad del modelo para cada estado nutricional. La sensibilidad indica la proporción de *casos positivos* correctamente identificados, respecto del total de casos positivos. La especificidad corresponde a la proporción de *casos negativos* clasificados correctamente, respecto del total de negativos. Se considera casos positivos a la categoría del estado nutricional que se calcula las métricas, y casos negativos las categorías restantes.

Cuadro 4.3.3: Tabla de clasificación del modelo logístico multinomial para el estado nutricional.

Observado	Pronosticado			Porcentaje correcto
	Normal o delgado	Sobrepeso	Obeso	
Normal o delgado	243	365	145	32,3 %
Sobrepeso	168	739	398	56,6 %
Obeso	91	458	694	55,8 %
Porcentaje global	15,2 %	47,3 %	37,5 %	50,8 %

Fuente: Elaboración propia.

La sensibilidad y especificidad para el estado nutricional normal o delgado es del 32,3 % y 89,8 % respectivamente. El estado nutricional sobrepeso tiene un 56,6 % de sensibilidad y 58,8 % de especificidad. Los valores para estado nutricional obeso son de 55,8 % y 73,6 % de sensibilidad y especificidad, respectivamente. En cuanto a la precisión se obtuvo 48,4 %, 47,3 %, y 56,1 % de precisión para los estados nutricionales normal o bajo peso, sobrepeso y obeso respectivamente.

Capítulo 5

Conclusión

En consideración con los resultados de la III Encuesta Nacional de Salud, el 74,2% de la población chilena presenta exceso de peso. Esta cifra sitúa a Chile como uno de los países con más altas tasas de sobrepeso y obesidad dentro de los países que pertenecen a la Organización para la Cooperación y el Desarrollo Económico. Los resultados que arroja esta investigación, en un futuro serán importantes a la hora de definir políticas públicas enfocadas en el área de salud, y en la implementación de intervenciones contra el sobrepeso y obesidad en Chile, ya que permiten guiar respecto a las características de las personas que presentan una mayor probabilidad de tener exceso de peso.

A pesar de que la literatura habla principalmente de la actividad física y alimentación como causas del sobrepeso y obesidad, en este estudio se distinguen otras características importantes en la población con exceso de peso, como el sexo, estado civil, zona geográfica donde se reside y nivel educacional, también se distingue variables conductuales como el consumo de alcohol y cigarrillos, consumo de agua y hábitos de sueño, y variables psicosociales como estrés y estrés financiero.

En este sentido, las personas que tienen mayor probabilidad de presentar sobrepeso en general están casados o en pareja, tienen entre 35 a 64 años,

beben alcohol 1 vez al mes, han dejado de consumir cigarrillos, presentan estrés financiero, padecen hipertensión y/o diabetes y tienen niveles alterados de colesterol LDL, triglicéridos y glucosa. Por otro lado, las personas que tienen una mayor probabilidad de presentar obesidad en general son mujeres casadas o en pareja, que residen en la zona norte o centro del país, su nivel educacional es bajo, beben alcohol entre 1 a 4 veces al mes, han dejado de consumir cigarrillos, duermen entre 1 a 5 horas diarias, se sienten estresados financieramente, padecen hipertensión y/o diabetes, y sus niveles de colesterol LDL, triglicéridos y glucosa son elevados en comparación con los niveles normales.

Mencionamos que la promoción de la salud debe ser de forma integral, promoviendo la actividad física, la prevención del consumo de alcohol y cigarrillos, y la buena alimentación. En la alimentación es donde se han enfocado las políticas públicas actuales en Chile, con la ley de etiquetado de alimentos, el programa “Elige vivir sano”, restricciones de publicidad, e impuestos al azúcar, a pesar de estas medidas el enfoque en la alimentación saludable sigue siendo complejo, dado que es difícil controlar la ingesta de alimentos con poca cantidad de nutrientes y alto contenido de grasa (comida chatarra), las familias con ingresos económicos menores están limitadas a la hora de seleccionar alimentos, y por la poca educación alimentaria y nutricional en niños y adultos a nivel nacional.

Es fundamental realizar de manera constante encuestas poblacionales en el área de la salud, ya que estas ayudan a distinguir las enfermedades que presenta la población, y como se ha avanzado con respecto a las políticas públicas aplicadas en los periodos anteriores. Es importante definir nuevas estrategias que permitan la prevención del sobrepeso y obesidad, y así en un futuro mejorar la calidad de vida de las personas residentes en Chile.

Por último, este trabajo despliega nuevas líneas de investigación en la determinación de los factores que inciden en la enfermedad del sobrepeso y obesidad en Chile. En primer lugar utilizando de manera combinada distintas herramientas de machine learning (como máquinas de soporte vectorial, potenciación del gradiente, redes neuronales), con distintos modelos (como regresión logística multinomial ordinal, modelo lineal generalizado mixto, o un modelo de selección de muestra de Heckman). Y en segundo lugar, empleando todas las encuestas nacionales de salud aplicadas en Chile hasta la fecha.





Anexo A

Descripción módulos

Cuadro A.0.1: Descripción módulos formulario 1 y 2 de la III Encuesta Nacional de Salud.

Módulo	Descripción
I Caracterización del entrevistado	Nacionalidad, idioma, residencia en Chile, pertenencia a pueblos indígenas.
II Evaluación cognitiva	Módulo aplicado solo a personas con 60 años o más. Preguntas relacionadas con la memoria y realización de actividades cognitivas con registro en el formulario.
III Calidad de vida y discapacidad	Preguntas relacionadas con el bienestar general del encuestado, su percepción de salud y dificultades para realizar actividades de la vida diaria.
IV Actividad física	Se pregunta sobre el tiempo que pasa realizando diferentes tipos de actividad física, como durante el trabajo, en el traslado de un lugar a otro o en el tiempo libre.
V Cardiovascular	Preguntas sobre síntomas de un accidente vascular o trombosis cerebral, tratamientos recibidos, o si ha sido diagnosticado de estas enfermedades.
VI Visión	Percepción de su vista, uso de lentes, operaciones y diagnósticos.
VII Audición	Percepción de la audición.
VIII Síntomas músculo esqueléticos	Preguntas que se refieren a los huesos, músculos, articulaciones o coyunturas.

(Continúa en la página siguiente)

Módulo	Descripción
IX Síntomas depresivos	Preguntas sobre el estado de ánimo, síntomas depresivos, pérdida de peso, problemas de concentración, diagnóstico y tratamiento de la depresión.
X Osteoporosis	Módulo aplicado solo a personas con 20 años o más. Preguntas relacionadas con fracturas de huesos.
XI Estado nutricional	Diagnóstico de exceso de peso y que está realizando para reducir el exceso de peso.
XII Dieta	Preguntas sobre la frecuencia del consumo de agua, pescado, lácteos, legumbres, frutas y verduras.
XIII Sueño	Cantidad de horas que duerme durante la semana y fin de semana, consumo de medicinas para dormir y problemas para conciliar el sueño.
XIV Tabaquismo	Preguntas relacionadas con el consumo de tabaco, edad que empezó a fumar, cantidad de cigarrillos, ha intentado dejar de fumar, entre otras.
XV Salud bucal	Percepción de salud bucal, frecuencia de visita al dentista y los motivos de visita al profesional.
XVI Encuesta de hipertensión	Diagnóstico de presión alta, tratamientos para mantener controlada la presión, tipo de tratamiento y tipo de establecimiento de salud donde se controla.
XVII Diabetes	Diagnóstico de diabetes, edad del diagnóstico, tipo de tratamiento y tipo de establecimiento donde se controla habitualmente.
XVIII Dislipidemias	Preguntas relacionadas con el padecimiento de colesterol alto, como cuando fue la última vez que se controló, si lo han diagnosticado, edad del diagnóstico, si está recibiendo tratamiento y de que tipo.
XIX Antecedentes familiares	Preguntas sobre enfermedades de familiares, ya sea mamá, papá, o hermanos.
XX Psicosocial	Preguntas sobre algunos pensamientos del entrevistado y si tiene apoyo en distintas situaciones propuestas en la entrevista, y preguntas sobre síntomas de estrés o angustia.

(Continúa en la página siguiente)

Módulo	Descripción
XXI Nivel socioeconómico del hogar	Estado civil, sistema previsional, nivel educacional más alto alcanzado, situación laboral, nivel socioeconómico y tipo de vivienda (agua, luz, eliminación de excretas y calefacción).
III Examen de sangre (Formulario 2)	Se realiza la toma de muestra de sangre.
VII Consumo de alcohol (Formulario 2)	Preguntas relacionadas con el consumo de alcohol o bebidas con alcohol, cantidad y frecuencia del consumo.

Fuente: Elaboración propia en base a formularios de la III Encuesta Nacional de Salud 2016-2017 (EPI, 2021).



Anexo B

Resumen gráfico de variables

Figura B.0.1: Porcentaje del estado nutricional dentro de la variable sexo.



Figura B.0.2: Porcentaje del estado nutricional dentro de la variable estado civil.

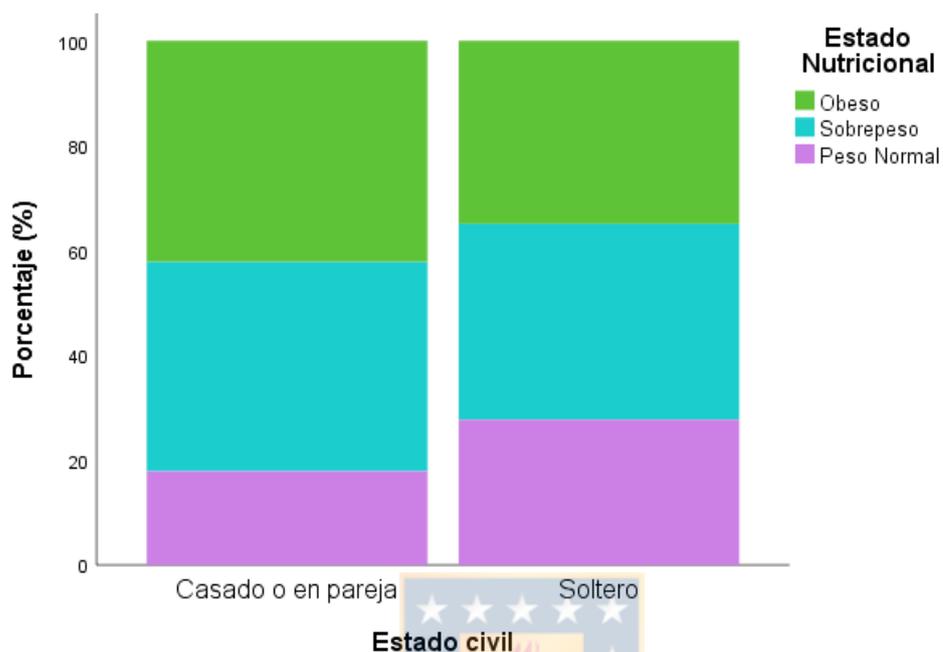


Figura B.0.3: Porcentaje del estado nutricional dentro de la variable sexo x estado civil.

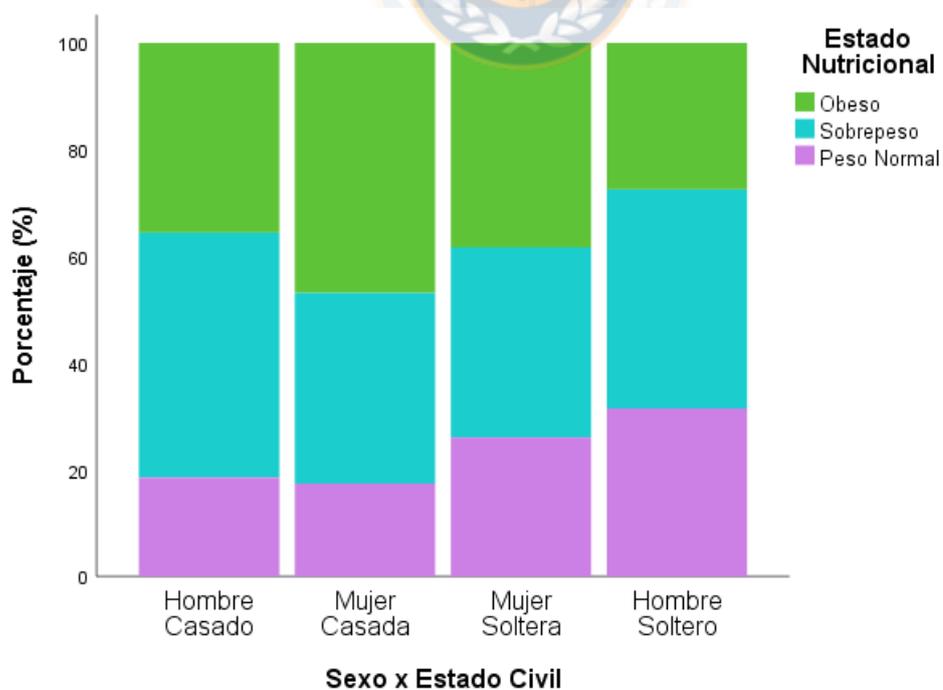


Figura B.0.4: Porcentaje del estado nutricional dentro de la variable edad.

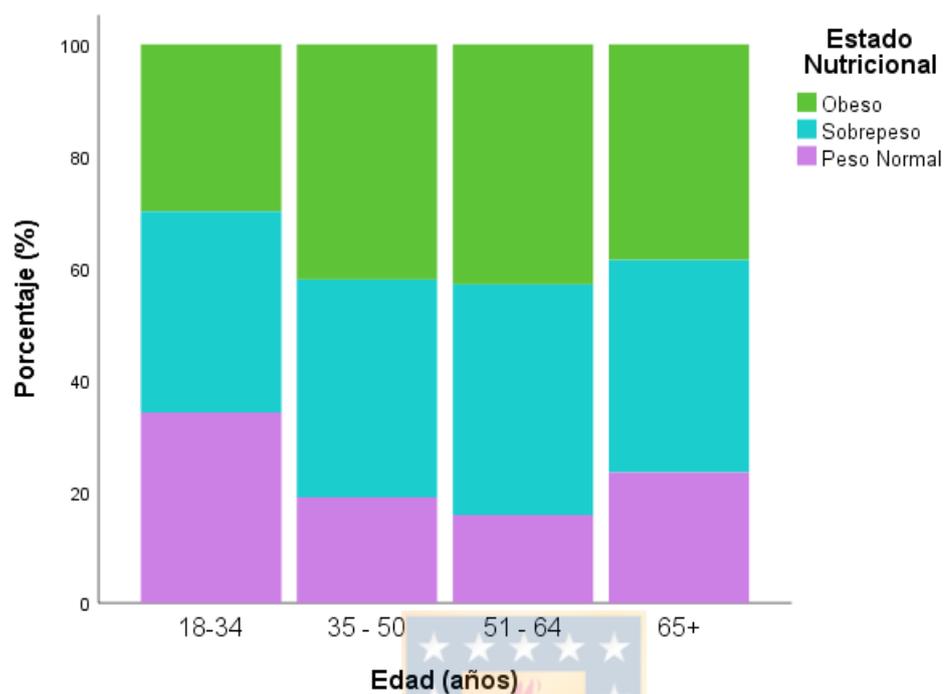


Figura B.0.5: Porcentaje del estado nutricional dentro de la variable zona geográfica.

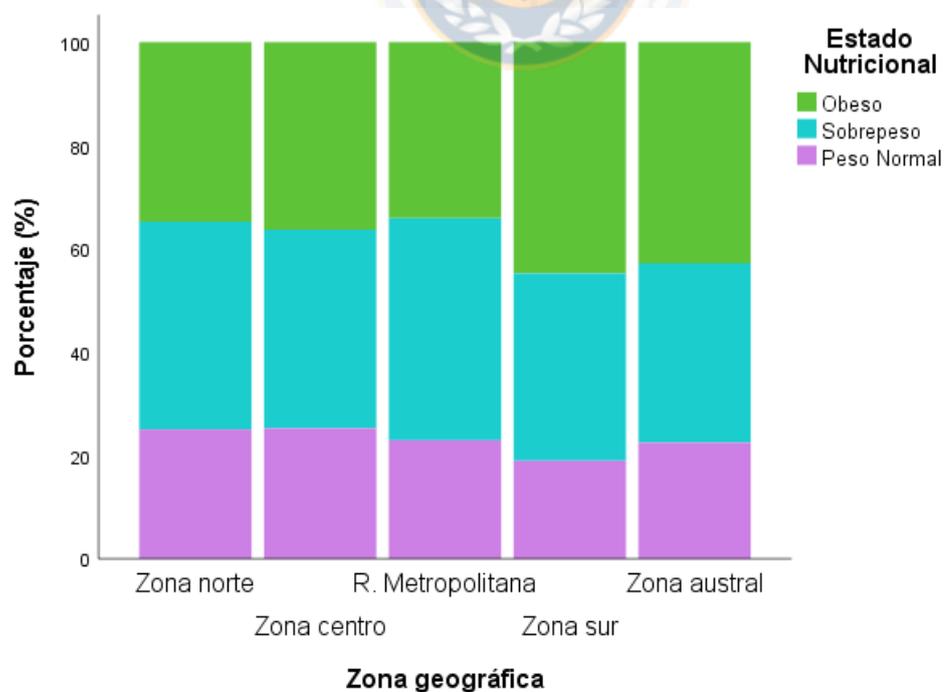


Figura B.0.6: Porcentaje del estado nutricional dentro de la variable área geográfica.

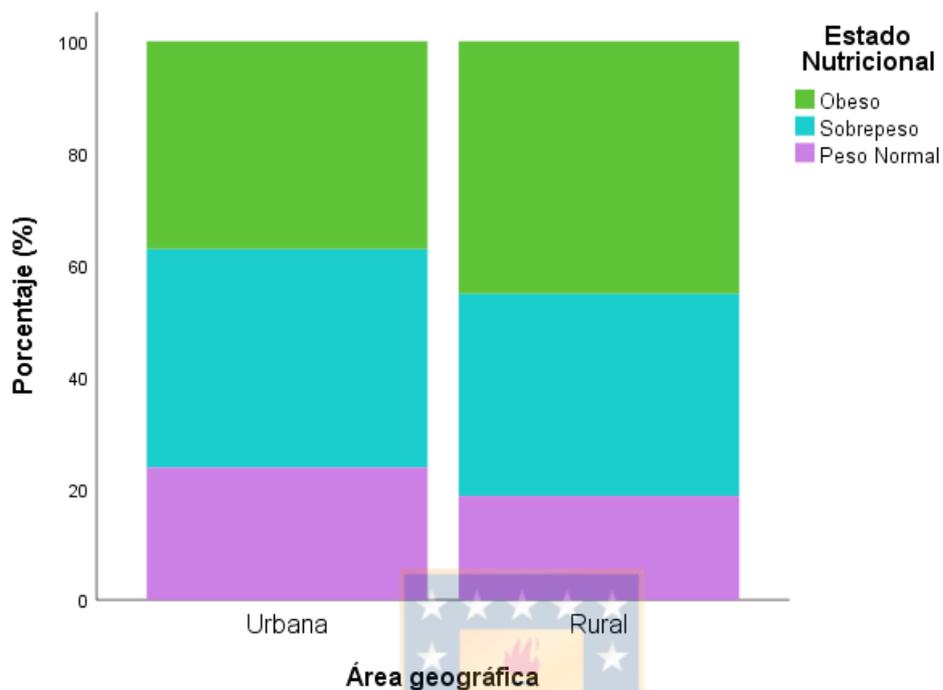


Figura B.0.7: Porcentaje del estado nutricional dentro de la variable nivel educacional.

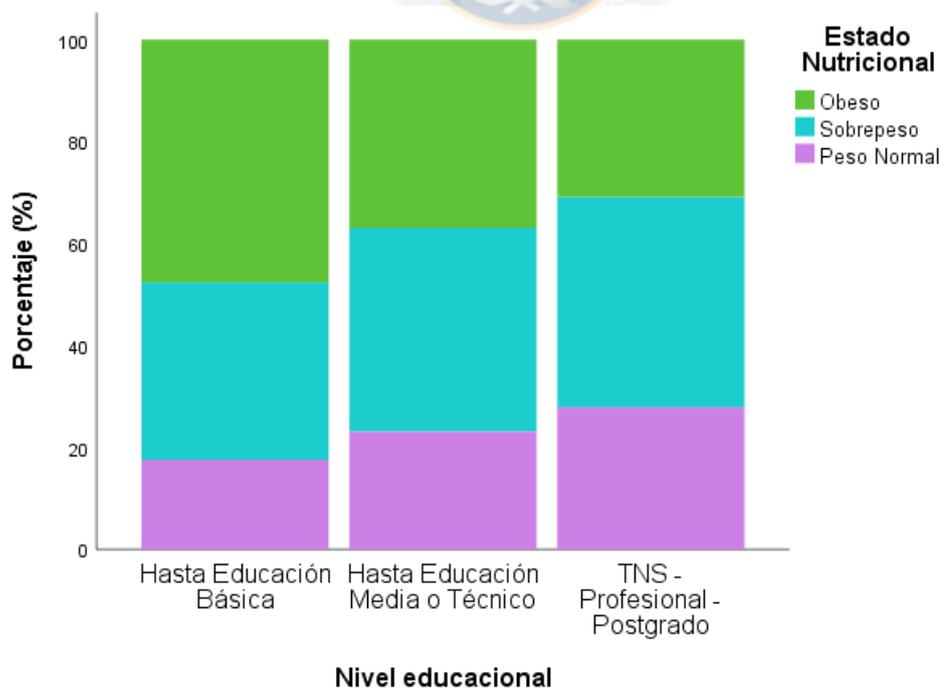


Figura B.0.8: Porcentaje del estado nutricional dentro de la variable consumo de alcohol.

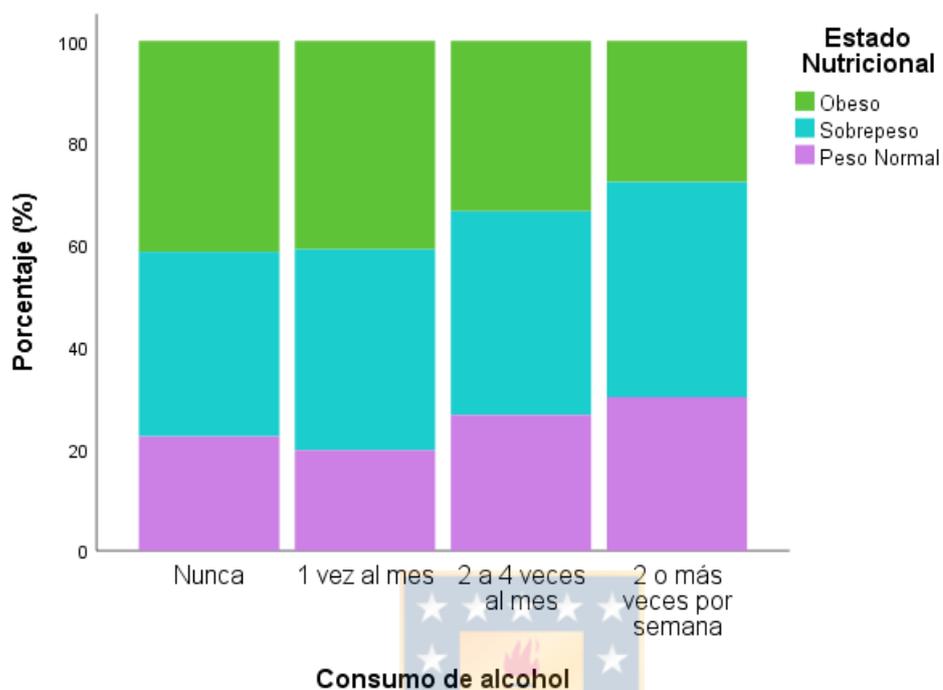


Figura B.0.9: Porcentaje del estado nutricional dentro de la variable consumo de cigarrillos.

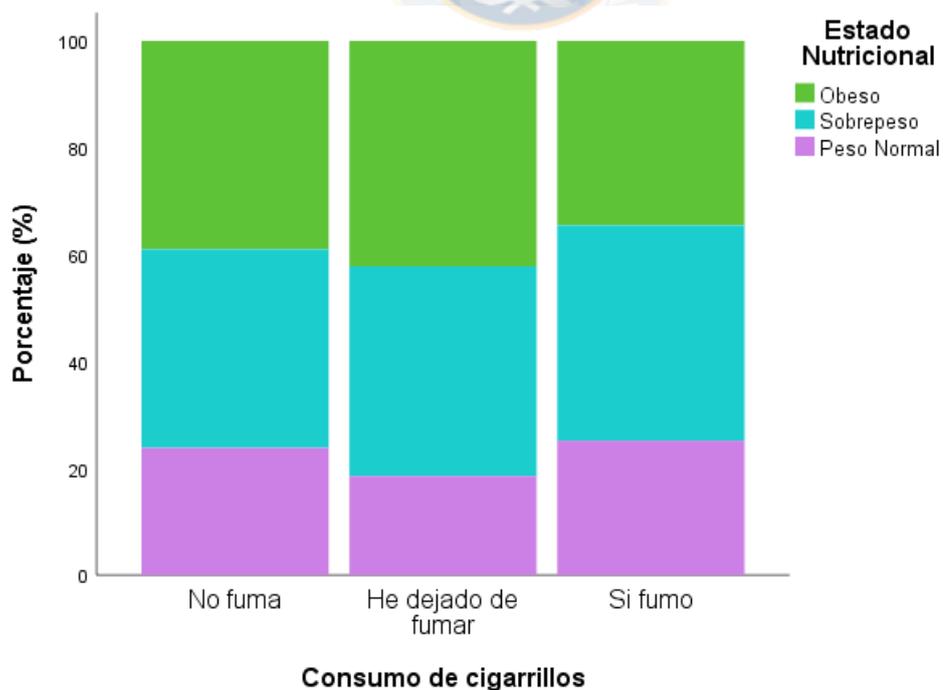


Figura B.0.10: Porcentaje del estado nutricional dentro de la variable dieta saludable.

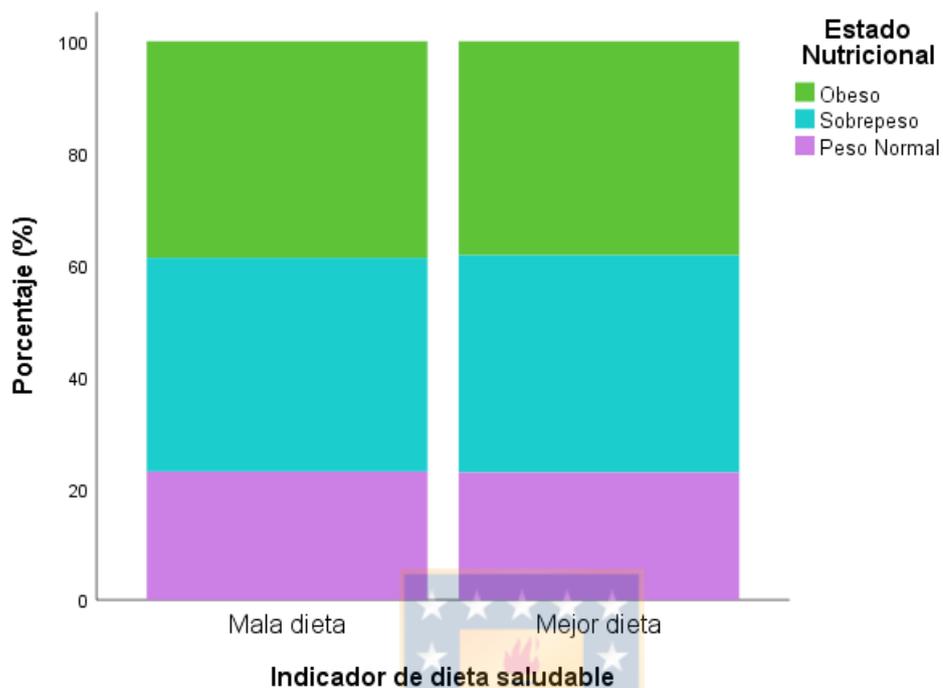


Figura B.0.11: Porcentaje del estado nutricional dentro de la variable actividad física.

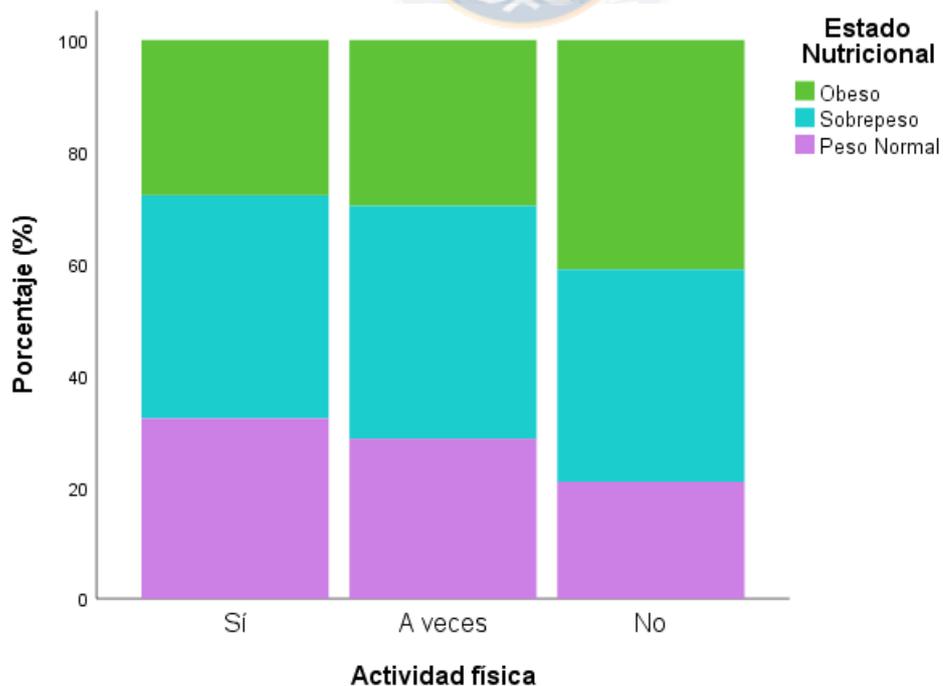
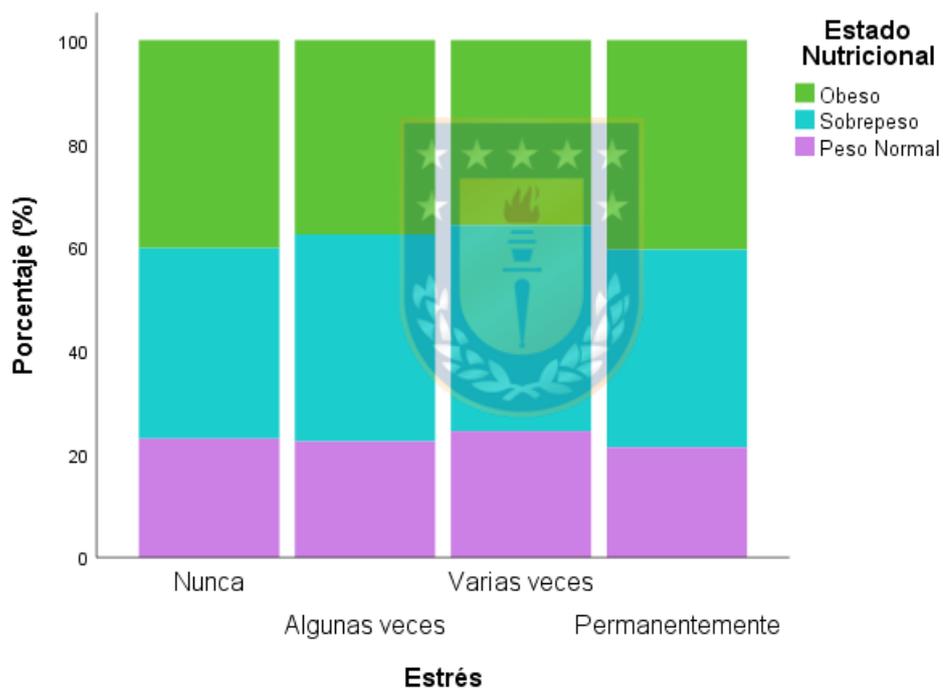


Figura B.0.12: Porcentaje del estado nutricional dentro de la variable estrés.



Referencias

- Afshin, A., Forouzanfar, M. H., Reitsma, M. B., Sur, P., Estep, K., y Lee, A. ... Murray, C. J. L. (2017). *Health Effects of Overweight and Obesity in 195 Countries over 25 Years*. *New England Journal of Medicine*, 377(1):13–27. <https://doi.org/10.1056/NEJMoa1614362>.
- Atalah, S. E. (2012). *Epidemiología de la obesidad en Chile*. *Revista Médica Clínica Las Condes*, 23(2):117–123. [https://doi.org/10.1016/S0716-8640\(12\)70287-0](https://doi.org/10.1016/S0716-8640(12)70287-0).
- Bennett, G., Wolin, K., y Duncan, D. (2008). Social Determinants of Obesity In FB. Hu (Ed.), *Obesity Epidemiology: Methods and Applications*, pp. 342–376. Oxford University Press.
- Castro, M., Muros, J., Cofré, C., Zurita, F., Chacón, R., y Espejo, T. (2018). *Índices de sobrepeso y obesidad en escolares de Santiago (Chile)*. *Journal of Sport and Health Research*, 10(2):251–256.
- Cerda, R. (2018). *Informe de Finanzas Públicas Ley de Presupuestos 2019*. https://www.dipres.gob.cl/597/articles-178468_a_presentacion_IFP_2019.pdf.
- Departamento de Epidemiología [EPI] (2017). *Encuesta Nacional de Salud 2016-2017. Primeros resultados*. https://www.minsal.cl/wp-content/uploads/2017/11/ENS-2016-17_PRIMEROS-RESULTADOS.pdf.
- Departamento de Epidemiología [EPI] (2021). *ENS - Encuesta Nacional de Salud*. Recuperado en mayo de 2020, de: <http://epi.minsal.cl/encuesta-ens/>.
- Ertel, W. (2017). *Undergraduate Topics in Computer Science, Introduction to Artificial Intelligence, Second Edition*. Springer.
- Gundersen, C., Mahatmya, D., Garasky, S., y Lohman, B. (2011). *Linking psychosocial stressors and childhood obesity*. *Obesity Reviews*, 12(5):e54–e63. <https://doi.org/10.1111/j.1467-789X.2010.00813.x>.
- Igual, L. y Seguí, S. (2017). *Introduction to Data Science: A Python Approach to Concepts, Techniques and Applications*. Springer Cham. <https://doi.org/10.1007/978-3-319-50017-1>.
- Junta Nacional de Auxilio Escolar y Becas [JUNAEB] (2019). *Mapa Nutricional 2019*. <https://www.junaeb.cl/wp-content/uploads/2013/03/Mapa-Nutricional-2019-1.pdf>.
- Llewellyn, A., Simmonds, M., Owen, C. G., y Woolacott, N. (2016). *Childhood obesity as a predictor of morbidity in adulthood: a systematic review and meta-analysis*. *Obesity Reviews*, 17(1):56–67. <https://doi.org/10.1111/obr.12316>.

- López-Roldán, P. y Fachelli, S. (2016). *Análisis de regresión logística. Metodología de la Investigación social cuantitativa*. . Bellaterra (Cerdanyola del Vallès): Dipòsit Digital de Documents, Universitat Autònoma de Barcelona. Recuperado en septiembre de 2020, de: <https://ddd.uab.cat/record/163570>.
- Ministerio de Salud [MINSAL] (2006). *Estrategia global contra la obesidad (EGO Chile): propuesta de trabajo*. <http://bibliotecaminsal-chile.bvsalud.org/lildbi/docsonline/get.php?id=1189>.
- Ministerio de Salud [MINSAL] (2007). *Programa para adultos obesos o con sobrepeso, prediabéticos y/o prehipertensos*. <http://bibliotecaminsal-chile.bvsalud.org/lildbi/docsonline/get.php?id=3769>.
- Ministerio de Salud [MINSAL] (2011). *Tomo I: Encuesta Nacional de Salud ENS Chile 2009-2010*. <https://www.minsal.cl/portal/url/item/bcb03d7bc28b64dfe040010165012d23.pdf>.
- National Heart Lung and Blood Institute [NIH] (2021). *Overweight and Obesity*. Recuperado en diciembre de 2021, de: <https://www.nhlbi.nih.gov/health-topics/overweight-and-obesity>.
- Organization for Economic Co-operation and Development [OECD] (2019a). *Health at a Glance 2019: OECD Indicators*. OECD Publishing, Paris. <https://doi.org/10.1787/4dd50c09-en>.
- Organization for Economic Co-operation and Development [OECD] (2019b). *The Heavy Burden of Obesity: The Economics of Prevention*, OECD Health Policy Studies. OECD Publishing, Paris. <https://doi.org/10.1787/67450d67-en>.
- Organization for Economic Co-operation and Development [OECD] (2019c). *OECD Reviews of Public Health: Chile. A Healthier Tomorrow*, OECD Reviews of Public Health. OECD Publishing, Paris. <https://doi.org/10.1787/9789264309593-en>.
- Reilly, J. J., Methven, E., McDowell, Z. C., Hacking, B., Alexander, D., Stewart, L., y Kelnar, C. J. H. (2003). *Health consequences of obesity*. Archives of Disease in Childhood, 88(9):748–752. <https://doi.org/10.1136/adc.88.9.748>.
- Romo-Muñoz, R., Dote-Pardo, J., Garrido-Henríquez, H., Araneda-Flores, J., y Gil, J. (2020). *Blueberry consumption and healthy lifestyles in an emerging market*. Spanish Journal of Agricultural Research, 17(4). <https://doi.org/10.5424/sjar/2019174-14195>.
- Schubert, M. N. y Ávalos, D. (2020). *Sistemas alimentares globais e lei de rotulagem de alimentos no Chile*. Redes, 25(2):527–544. <https://doi.org/10.17058/redes.v25i2.14850>.
- Singh, A. S., Mulder, C., Twisk, J. W. R., Van Mechelen, W., y Chinapaw, M. J. M. (2008). *Tracking of childhood overweight into adulthood: a systematic review of the literature*. Obesity Reviews, 9(5):474–488. <https://doi.org/10.1111/j.1467-789X.2008.00475.x>.
- Vielma, M., Vicente, B., Rioseco, P., Castro, N., y Torres, S. (2002). *Validación en Chile de la entrevista diagnóstica estandarizada para estudios epidemiológicos "CIDI"*. Revista de psiquiatría 1992, 9(1):1039–1049.

- Vio, F., Albala, C., y Kain, J. (2008). *Nutrition transition in Chile revisited: mid-term evaluation of obesity goals for the period 2000–2010*. *Public Health Nutrition*, 11(4):405–412. <https://doi.org/10.1017/S136898000700050X>.
- Vio del Rio, F. (2018). *Aumento de la obesidad en Chile y en el mundo*. *Revista chilena de nutrición*, 45(1):6. <http://dx.doi.org/10.4067/s0717-75182018000100006>.
- World Health Organization [WHO] (2000). *Informe sobre la salud en el mundo 2000. Capítulo 1: ¿Por qué son importantes los sistemas de salud?*. https://www.who.int/whr/2000/en/whr00_ch1_es.pdf.
- World Health Organization [WHO] (2020). *Body mass index - BMI*. Recuperado en diciembre de 2020, de: <https://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi>.
- World Health Organization [WHO] (2021). *Obesity and overweight*. <https://www.who.int/es/news-room/fact-sheets/detail/obesity-and-overweight>.
- Xu, S. y Xue, Y. (2016). *Pediatric obesity: Causes, symptoms, prevention and treatment (Review)*. *Experimental and Therapeutic Medicine*, 11(1):15–20. <https://doi.org/10.3892/etm.2015.2853>.

