

Taller 1 R

Silvana Ferro - 202320772

Gabriela Fernandez

Catalina Barbosa moreno 202310281

Paula Andrea Bernal Mongui- 202423557

1. ¿El tipo de estudio que representa la ENSIN puede demostrar causalidad? (Describa el tipo de estudio y justifique su respuesta)

No, la ENSIN no puede demostrar causalidad porque es un estudio transversal o de prevalencia.

La ENSIN 2010 es un estudio observacional transversal. Esto significa que:

- Observacional: Los investigadores solo miden y observan variables (como el estado nutricional, el nivel de ingresos, o los hábitos alimenticios) sin intervenir o modificar las condiciones de los participantes.
- Transversal: Todos los datos se recogen en un único punto en tiempo. Es como tomar una fotografía de la población colombiana en el año 2010 para ver cómo era su situación nutricional en ese momento específico.

La causalidad requiere demostrar tres cosas:

1. que la causa ocurre antes que el efecto,
2. que existe una asociación entre ellas, y
3. que se han descartado otras explicaciones alternativas (factores de confusión).

Para demostrar causalidad, un estudio debe tener estas características esenciales:

1. Secuencia temporal (Temporalidad): La causa potencial debe ocurrir antes que el efecto. Este es el requisito fundamental y absoluto.
2. Asociación: Debe existir una relación estadística entre la causa potencial y el efecto. Un cambio en la causa debe estar vinculado a un cambio en el efecto.
3. Eliminación de explicaciones alternativas (No confusión): Se deben descartar o controlar todos los factores de confusión (otras variables que podrían ser la causa real de la asociación observada).

2. ¿Qué características debe tener un estudio para poder demostrar causalidad?

1. Secuencia temporal: La causa potencial debe medirse antes de que ocurra el efecto. Esto solo se logra con estudios longitudinales que siguen a los participantes en el tiempo.
2. Manipulación o comparación controlada: Idealmente, el estudio debe ser experimental. Esto significa que los investigadores asignan al azar a los participantes a un grupo que recibe la intervención (la causa potencial) y a un grupo de control que no la recibe (como en un ensayo clínico aleatorizado). Esto permite aislar el efecto de la causa.

3. Control de factores de confusión: El diseño debe permitir identificar y eliminar la influencia de otras variables que podrían falsear la relación (como la edad, el nivel socioeconómico o la genética). La aleatorización es la mejor herramienta para esto.
4. Asociación fuerte y consistente: Debe existir un vínculo estadístico claro y robusto entre la causa y el efecto, donde un cambio en la primera produzca un cambio predecible en el segundo (relación dosis-respuesta).

Solo los diseños experimentales (como los ensayos clínicos) o los observacionales longitudinales muy robustos (estudios de cohorte) pueden aspirar a demostrar causalidad, porque establecen el orden de los eventos y controlan las variables alternativas. Los estudios transversales, como la ENSIN, no cumplen estos requisitos.

3. ¿Cuál es la muestra de la población a estudio de la ENSIN para el análisis de sobrepeso y obesidad?

Población de estudio:

Personas de 0 a 64 años residentes en hogares colombianos.

Tamaño de la muestra para antropometría (sobrepeso/obesidad):

- 162,331 personas fueron medidas (peso, talla, circunferencia de cintura) en todo el territorio nacional.
- Esta muestra es representativa a nivel nacional, urbano/rural, para 6 regiones, 14 subregiones y 32 departamentos.

Desglose por grupos de edad:

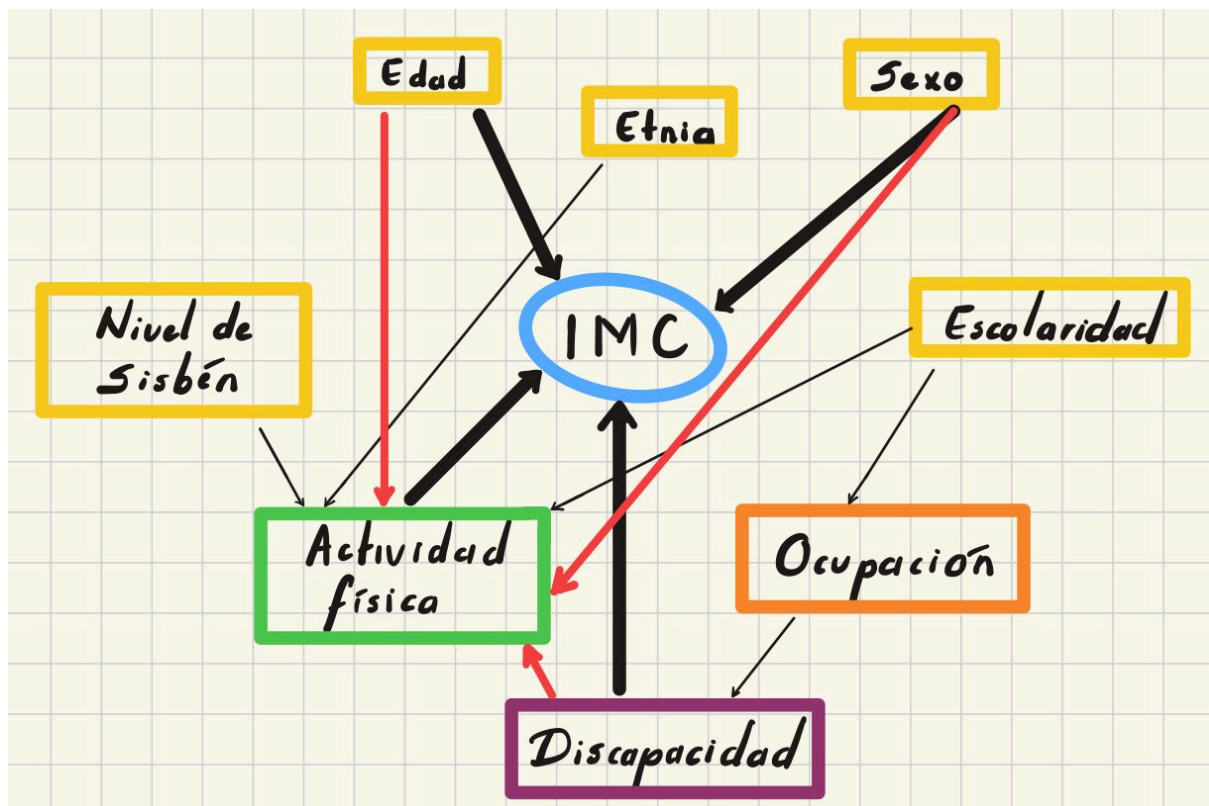
1. Menores de 5 años: Incluidos en la muestra total de antropometría.
2. Niños y adolescentes de 5 a 17 años: Analizados para sobrepeso/obesidad usando estándares de la OMS.
3. Adultos de 18 a 64 años: Clasificados con puntos de corte de IMC y circunferencia de cintura (IDF).
4. Gestantes: Evaluadas con criterios de IMC gestacional (Atalah et al.), aunque esta submuestra no es representativa a nivel nacional.

Características clave de la muestra:

- Cobertura: 85% de la población objetivo (0-64 años).
- Estratificación: Representativa por sexo, edad, etnia, nivel socioeconómico (SISBEN), y zona (urbana/rural).
- Exclusión: No incluye población institucionalizada (ej: hogares geriátricos, bases militares).

La ENSIN 2010 utilizó una muestra robusta de 162,331 personas para estimar la prevalencia de sobrepeso y obesidad en Colombia, con representatividad estadística para múltiples subgrupos poblacionales y regiones.

4. Entregar el gráfico del modelo conceptual del análisis que estamos realizando, teniendo en cuenta sólo las variables de la base de datos. (Tenga en cuenta el ejemplo de la doctora en clase).



5. Justificar TODAS las relaciones representadas en el modelo conceptual (la justificación incluye la dirección, el sentido y la magnitud de las asociaciones descritas en el modelo) (esta justificación no debe ser más larga de 200 palabras)

El IMC está como variable final, influido por determinantes biológicos y sociales. Las relaciones más fuertes son Edad hacia IMC, Sexo hacia IMC, Discapacidad hacia IMC y Actividad física hacia IMC: la edad incrementa el IMC disminuyendo el metabolismo y cambiando la composición corporal; el sexo condiciona diferencias hormonales y metabólicas en la repartición de grasa; la discapacidad restringe la movilidad e incrementa el riesgo de obesidad; La actividad física modula directamente el balance energético. Con un grosor intermedio están Edad hacia Actividad física, Sexo hacia Actividad física y Discapacidad hacia Actividad física, porque estas variables afectan el nivel de ejercicio realizado, pero su influencia sobre el IMC se da de manera mediada. Finalmente, las flechas más delgadas representan asociaciones indirectas: Escolaridad hacia Actividad física, pues la educación facilita hábitos saludables; Escolaridad hacia Ocupación, dado que el nivel académico determina oportunidades laborales; Ocupación hacia Discapacidad, ya que en algunos trabajos favorecen riesgos físicos o limitaciones; Etnia hacia Actividad física, vinculada a tradiciones culturales y estilos de vida; y Nivel de Sisbén hacia Actividad física, reflejando desigualdades en acceso a recursos y espacios para el ejercicio. En conjunto, el modelo diferencia claramente influencias directas, mediadas e indirectas sobre el IMC.

6. Escribir la pregunta de estudio del análisis que estamos realizando (Pregunta de investigación).

- ¿Cuál es la asociación entre los factores sociodemográficos (edad, sexo, nivel educativo, ocupación) y la presencia de sobrepeso u obesidad en la población colombiana según los datos de la ENSIN?

7. Escribir la hipótesis nula y alterna del estudio y análisis que estamos realizando.

- **La hipótesis nula:** establece que no existe asociación estadísticamente significativa entre los factores sociodemográficos (edad, sexo, nivel educativo y ocupación) y la presencia de sobrepeso u obesidad en la población colombiana según los datos de la ENSIN.
- **La hipótesis alterna:** plantea que sí existe una asociación estadísticamente significativa entre al menos uno de los factores sociodemográficos mencionados y el sobrepeso u obesidad en la población colombiana.
- El análisis descriptivo de las variables de interés se realiza mediante frecuencias absolutas, relativas y medidas de tendencia central. Así mismo, se aplican diferentes pruebas para evaluar la relación entre variables categóricas y, de esta manera, se llevará a cabo una regresión logística binaria, donde la variable dependiente será la presencia de sobrepeso/obesidad y las variables independientes serán la edad, el sexo, el nivel educativo y la ocupación.

8. Describa el tipo de las variables que se están analizando (Tipo de numérica, tipo de categórica).

- **Edad (EDAD / EDADNEW)**
 - Original (EDAD): **Numérica continua** (años, 18–64).
 - Nueva (EDADNEW): **Categórica ordinal** (1 = 18–25; 2 = >25).
- **Sexo (SEXO / SEXONEW)**
 - Original (SEXO): **Categórica nominal** (1 = Hombre; 2 = Mujer).
 - Nueva (SEXONEW): **Categórica nominal** (0 = Hombre; 1 = Mujer).
- **Etnia (ETNIA / ETNIANEW)**
 - Original (ETNIA): **Categórica nominal** (varios grupos étnicos).
 - Nueva (ETNIANEW): **Categórica nominal** (agrupación en 0 y 1 según pertenencia a minorías étnicas).

- **Escolaridad (ESCOLARIDAD / ESCOLARIDADNEW)**

- Original (ESCOLARIDAD): **Categórica ordinal** (niveles educativos desde ninguno hasta posgrado).
- Nueva (ESCOLARIDADNEW): **Categórica ordinal** (0 = básica/media; 1 = técnico, universidad o posgrado).

- **Ocupación (OCUPACION / OCUPACIONNEW)**

- Original (OCUPACION): **Categórica nominal** (trabajó, estudió, pensión, oficios, etc.).
- Nueva (OCUPACIONNEW): **Categórica nominal** (agrupación binaria: 0 = trabajó/tenía trabajo; 1 = otras condiciones).

- **Discapacidad (QH42 / DISCAPACIDADNEW)**

- Original (DISCAPACIDAD): **Categórica nominal** (sí/no).
- Nueva (DISCAPACIDADNEW): **Categórica nominal** (0 = No; 1 = Sí).

- **Nivel de Sisbén (SISBEN / SISBENNEW)**

- Original (SISBEN): **Categórica ordinal** (1, 2, 3, 4–6).
- Nueva (SISBENNEW): **Categórica ordinal** (0 = Sisbén 1–3; 1 = Sisbén 4–6).

- **Actividad física en tiempo libre (Tlmeett / TlmeettNEW)**

- Original (Tlmeett): **Categórica nominal dicotómica** (0 = no cumple; 1 = cumple recomendaciones OMS).
- Nueva (TlmeettNEW): **Categórica nominal dicotómica** (igual codificación).

- **Índice de Masa Corporal (IMC / imcNEW)**
- Original (IMC): **Numérica continua** (kg/m²).
- Nueva (imcNEW): **Categoría ordinal** (0 = delgadez; 1 = normal; 2 = sobrepeso; 3 = obesidad).

Esta clasificación nos permite diferenciar entre las mediciones directas obtenidas de la encuesta y las variables derivadas. Estas son fundamentales para el análisis epidemiológico de sobrepeso y obesidad.

9. Haga una tabla con los porcentajes y frecuencias de las variables como si fuera para publicación en el AJE. En la tabla recuerde la importancia del título y el nombre de las variables. Consulte el ejemplo de tabla de un artículo publicado en el AJE que está en BLOQUE NEÓN para el formato de tabla correcto.

Tabla 1. Características de la Población

Característica	Categoría	n	%
Edad — años	18–25	3291	22.8
	>25–64	11170	77.2
Sexo — n (%)	Hombre	6240	43.1
	Mujer	8221	56.9
Nivel SISBEN — n (%)	Niveles 1–3	8801	60.9
	Niveles 4–6	5660	39.1
Nivel educativo (máximo) — n (%)	Básico/Secundaria	11028	76.5
	Técnico/Universitario/Postgrado	3385	23.5
Condición laboral (última semana) — n (%)	Activo	8856	62.6
	Inactivo	5291	37.4
Discapacidad (QH42) — n (%)	Sí	852	5.9
	No	13609	94.1
Actividad física en tiempo libre (recomendación OMS) — n (%)	No cumple con la recomendación de la OMS	11524	79.7
	Cumple con la recomendación de la OMS	2937	20.3
Exceso de peso (IMC ≥25) — n (%)	No	6182	45.7
	Sí	7354	54.3

* Los porcentajes pueden no sumar 100% debido al redondeo. OMS = Organización Mundial de la Salud; IMC = Índice de Masa Corporal.

10. SÓLO en un párrafo, describa los resultados de la tabla de la pregunta anterior (p.9)

La muestra adulta (18–64 años) es 56,9% mujeres; 77,3% y tienen más de 25 años. El 60,9% pertenece a SISBEN (Estratos 1,2,3) y 24,4% reporta educación técnica/universitaria/posgrado. El 62,6% está activo laboralmente. El 5,9% reporta discapacidad. Solo 20,3% cumple recomendaciones de actividad física en el tiempo libre. Por último, el exceso de peso afecta aproximadamente a 1 de cada 2 adultos (54,3%).

11. Hacer una tabla con sólo los porcentajes de missings de cada variable nueva

Variable	% de faltantes (NA)
Exceso de peso	6,40 %
Ocupación	2,17 %
Escolaridad	0,33 %
Edad (nueva)	0,00 %
Sexo	0,00 %
Sisbén (niveles)	0,00 %
Discapacidad	0,00 %
TLMEETT (cumplimiento)	0,00 %

12. ¿Cree usted que el porcentaje de missings de alguna variable es significativo y puede alterar la validez de los resultados?

Sí. El 6,4 % de datos faltantes en EXCESO_PESO_NEW (IMC) si bien es moderado, puede mostrar un sesgo en las estimaciones si la ausencia de estos datos no es completamente al azar. Por ejemplo, cuando las personas que no tienen IMC muestran diferencias importantes en comparación al resto, se encuentra un posible sesgo con la ausencia de posibles valores de IMC considerablemente altos o bajos.

13. Describa la media, desviación estándar, mediana, moda, rango y rango Inter cuartil de las

Media: 37,61 años. Útil como medida de tendencia central para comparar grupos y como base de muchos métodos inferenciales (intervalos, pruebas t), aunque es sensible a asimetrías y valores atípicos.

Desviación estándar: 12,90 años. Cuantifica la dispersión alrededor de la media en las mismas unidades; se usa para estandarizar (z-scores) y en diferentes modelos estadísticos que asumen normalidad (ANOVA, t, regresión).

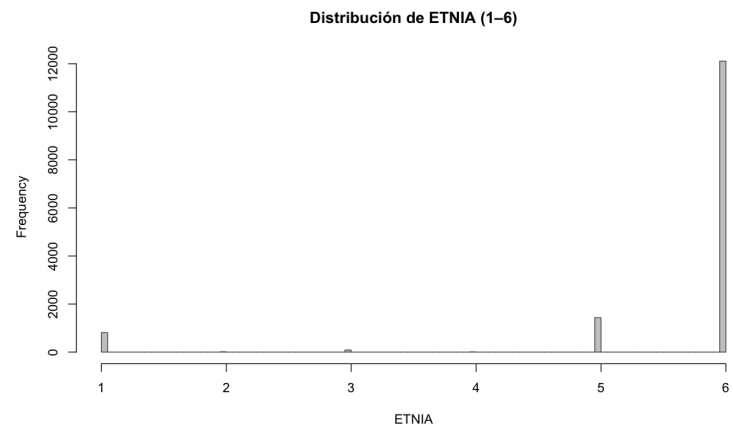
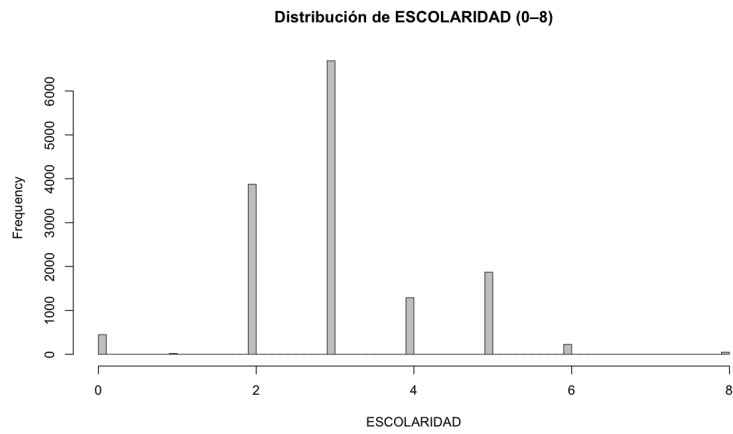
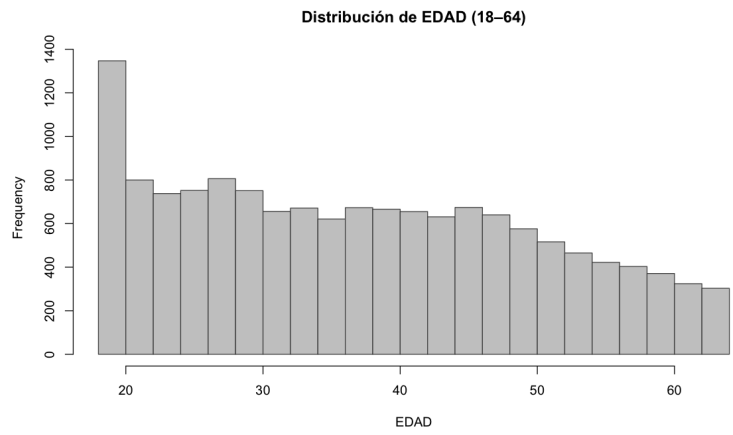
Mediana: 37 años. Punto central (50 % por debajo/encima), muy robusta frente a outliers y útil cuando la distribución es asimétrica.

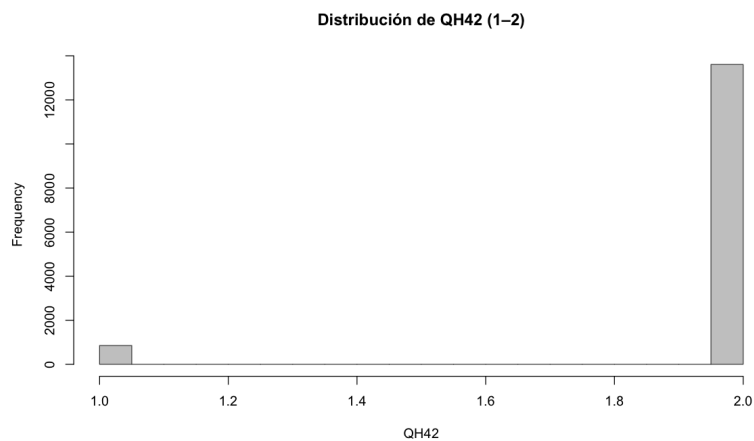
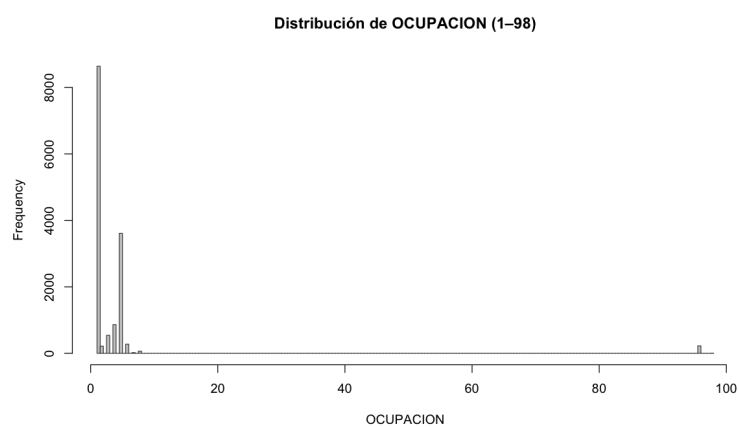
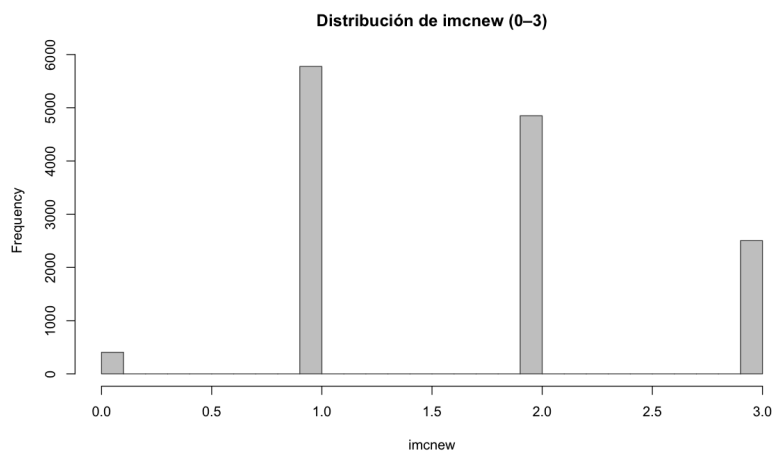
Moda: 18. Valor más frecuente, en variables continuas puede depender del ajuste polinomial que se aplique, pero ayuda a detectar multimodalidad y patrones de concentración.

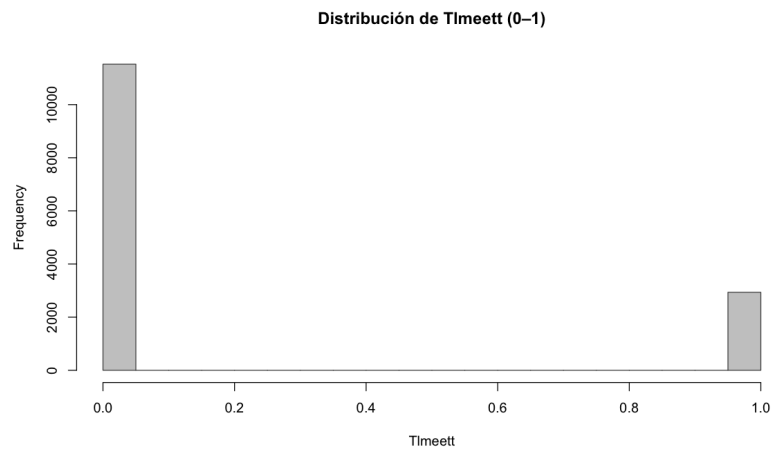
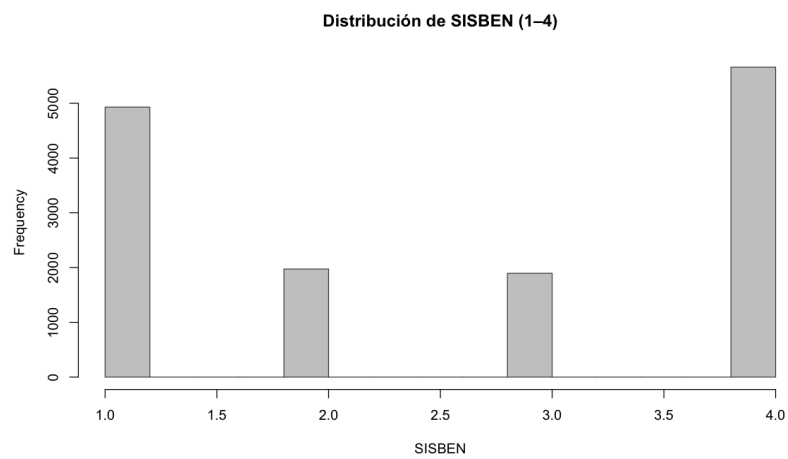
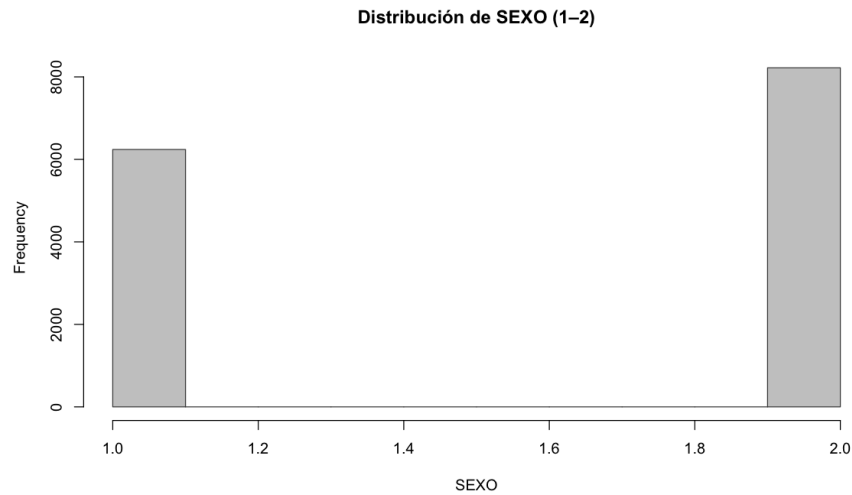
Rango: 18–64 años. Mide la extensión total (máx–mín) lo que lo vuelve muy sensible a outliers (datos atípicos) y al tamaño muestral, por lo que se usa como indicio más que como indicador normal.

Rango por cuartil (IQR): $Q3 - Q1 = 48 - 26 = 22$. Captura la variabilidad central del 50 % de los datos; es robusto a outliers y base de los diagramas de caja y bigote para comparar dispersión entre grupos.

14. Colocar las gráficas de la distribución de las variables numéricas y determinar qué tipo de distribución tienen. Analice y describa la kurtosis de las gráficas.







1. ETNIA: curtosis (exceso) = 9.8094, Leptocúrtica
2. OCUPACIÓN: curtosis (exceso) = 55.4461, Leptocúrtica.
3. QH42: curtosis (exceso) = 12.0356, Leptocúrtica.
4. EDAD: curtosis (exceso) = -1.0432, Platicúrtica
5. SEXO: curtosis (exceso) = -1.9235, Platicúrtica.
6. ESCOLARIDAD: curtosis (exceso) = 1.2832, Leptocúrtica.
7. SISBEN: curtosis (exceso) = -1.7237, Platicúrtica.
8. imcnew: curtosis (exceso) = -0.8436, Platicúrtica.

9. Tlmeett: curtosis (exceso) = 0.1786, Mesocúrtica

15. Comente acerca de la categorización de la variable escolaridad y ocupación. ¿Está usted de acuerdo o en desacuerdo? ¿Por qué?

Escolaridad_NEW (binaria): simplifica (básica/media vs superior). Si bien se pierde un importante gradiente en el análisis, existe una mayor población que cursa educación básica en conjunto con educación media y la mayor brecha es con la educación superior.

Ocupación_NEW (activo o inactivo): agrupa estados heterogéneos (estudio, hogar, jubilación) que pueden relacionarse distinto con AF/IMC. En general, es una categorización aceptable para este análisis, pero conviene probar categorías más finas para aumentar la sensibilidad.

16. ¿Qué categorización utilizó para IMC teniendo en cuenta que el análisis será enfocado a sobrepeso y obesidad? Justifique su respuesta

Se utilizó EXCESO_PESO_NEW donde se parametriza con: Si = sobrepeso u obesidad ($IMC \geq 25$) y No = delgadez/normal. Esto se hizo debido a que un índice de 25 como límite para el sobrepeso u obesidad es el indicador usado por ENSIN y se alinea con el objetivo de enfocarse en sobrepeso+obesidad como un único conjunto a desarrollar.

REFERENCIAS

- Kirkwood BR, Sterne JAC. Essential Medical Statistics. 2nd ed. Oxford: Blackwell Science; 2003.
- Ministerio de Salud y Protección Social de Colombia; Instituto Colombiano de Bienestar Familiar (ICBF). Encuesta Nacional de Situación Nutricional (ENSIN) 2015 [Internet]. Bogotá: MinSalud; 2017 [citado 2025 ago 29]. Disponible en: <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/ED/GCFI/ensin-colombia-2015.pdf>
-