

EBS

Encuesta de Bienestar Social

● ● ● ● ● 2025 ● ● ● ● ●

Metodología de Diseño Muestral

División Observatorio Social

Septiembre 2026



Serie Metodología Encuesta de Bienestar Social – EBS 2025

Metodología de Diseño Muestral

Septiembre de 2026.

Subdepartamento de Diseño de Marcos y Muestras

Departamento de Metodologías e Innovación Estadística

Instituto Nacional de Estadísticas

División Observatorio Social

Subsecretaría de Evaluación Social

Ministerio de Desarrollo Social y Familia

Informe editado por el Ministerio de Desarrollo Social y Familia, a partir del Informe de diseño muestral de la Encuesta de Bienestar Social, EBS 2025 elaborado por el Instituto Nacional de Estadísticas (INE).



Encuesta de Bienestar Social - EBS 2025

La Encuesta de Bienestar Social (EBS) es una iniciativa institucional del Ministerio de Desarrollo Social y Familia implementada por primera vez en el año 2021, cuyo objetivo es, en conjunto con la Encuesta de Caracterización Socioeconómica Nacional (Casen), entregar un diagnóstico completo del bienestar y de las oportunidades que tienen las personas en el país contribuyendo al diseño, desarrollo y evaluación de políticas públicas.

La EBS se aplica a una submuestra de personas entrevistadas previamente en Casen, lo que permite complementar la información socioeconómica levantada por dicha encuesta con antecedentes sobre percepciones, experiencias y evaluaciones subjetivas asociadas al bienestar. La población objetivo corresponde a personas de 18 años o más, cumplidos a la fecha de realización de Casen 2024, que residen en viviendas particulares en el territorio nacional. Esta vinculación metodológica permite analizar de manera integrada condiciones objetivas de vida, tales como trabajo, salud, vivienda o entorno, junto con dimensiones subjetivas referidas a la forma en que las personas evalúan su calidad de vida, sus relaciones, su seguridad, su uso del tiempo y otros aspectos relevantes de su bienestar.

El marco conceptual de la EBS se fundamenta en las directrices de bienestar de la Organización para la Cooperación y el Desarrollo Económicos (OCDE), particularmente en la serie metodológica: *How's Life?: Measuring Well-Being*¹, así como en el enfoque de capacidades de Amartya Sen. Desde la perspectiva de la OECD, el bienestar se entiende como un fenómeno multidimensional que considera tanto las condiciones objetivas de vida como la forma en que las personas las experimentan y evalúan. Por su parte, el enfoque de capacidades plantea que el bienestar no depende solo de los recursos disponible, sino también de las oportunidades reales que tienen las personas para desarrollar sus proyectos de vida y alcanzar aquello que valoran. En conjunto, ambos enfoques permiten analizar el bienestar desde una perspectiva amplia, considerando condiciones materiales, experiencias subjetivas y oportunidades efectivas.

En su tercera versión, correspondiente a la serie 2025, la EBS incorpora los lineamientos actualizados de los organismos internacionales, integrando de forma explícita las tres dimensiones del bienestar subjetivo propuestas en el informe *Guidelines on Measuring Subjective Well-being* (OCDE, 2025): evaluación de la vida, afectos y sentido o propósito.

El cuestionario se estructura en once módulos analíticos orientados a las siguientes dimensiones: bienestar subjetivo, educación, trabajo en la ocupación, uso del tiempo, ingresos, salud, vivienda, calidad del medio ambiente, seguridad, relaciones sociales; y confianza y participación.

El diseño muestral corresponde a un muestreo probabilístico y bifásico, cuya distribución garantiza representatividad estadística para el parámetro de interés que corresponde a la tasa de insatisfacción con la vida a niveles nacional, regional y por área geográfica urbano-rural.

El levantamiento se ejecutó entre el 23 de julio y el 12 de octubre de 2025, con modalidad telefónica mediante encuestadores asistidos por computador (CATI), utilizando como soporte tecnológico de captura el software Survey Solutions.

¹ Para más detalle ver: https://www.oecd.org/en/publications/how-s-life_23089679.html



Tanto las bases de datos analizadas como la documentación metodológica —incluyendo manuales, libros de códigos, cuestionarios y marcos de referencia de la EBS 2025— se encuentran disponibles de manera abierta en el repositorio oficial del Banco Integrado de Datos (BIDAT)².

²https://bidat.gob.cl/directorio/Encuestas%20Observatorio%20Social/encuesta-de-bienestar-social-2025?name_subnivel=EBS%20-%20Encuesta%20de%20Bienestar%20Social



ÍNDICE

1.	ANTECEDENTES.....	9
2.	CARACTERÍSTICAS GENERALES DEL DISEÑO MUESTRAL.....	10
2.1.	Objetivo.....	10
2.2.	Población objetivo	10
2.3.	Unidad de información	10
2.4.	Marco muestral.....	10
2.5.	Dominios de estudio	11
2.6.	Objetivo de precisión y tamaño de muestra	11
2.7.	Estrategia muestral.....	11
3.	ESTIMACIÓN DEL TAMAÑO MUESTRAL.....	12
3.1.	Etapas para la estimación del tamaño muestral para EBS 2025.....	13
3.2.	Parámetros utilizados para el cálculo del tamaño muestral.....	17
3.3.	Metodología de cálculo del tamaño muestral utilizando muestras complejas.....	18
3.4.	Resumen proceso de simulación	19
3.5.	Escenario escogido.....	22
3.5.1.	Escenario 1.3.....	22
3.6.	Tamaño muestral propuesto.....	23
3.7.	Errores esperados	25
4.	SELECCIÓN DE UNIDADES MUESTRALES.....	26
5.	CONSTRUCCIÓN FACTORES DE EXPANSIÓN	29
5.1.	Factor de expansión de primera fase.....	30
5.2.	Factor de expansión de segunda fase	31
5.2.1.	Ponderador de selección de personas como inverso de la probabilidad de selección.....	31
5.2.2.	Producto del factor de primera y segunda fase	32
5.3.	Ajustes por no contacto, elegibilidad y no respuesta.....	32
5.3.1.	Ajuste por no contacto	33
5.3.2.	Ajuste por no elegibilidad.....	35
5.3.3.	Ajuste por no respuesta.....	36
5.4.	Suavizamiento	38
5.4.1.	Método de contracción a la media (CM)	38
5.4.2.	Método R-K	39
5.4.3.	Método mixto	39
5.5.	Calibración de factores de expansión.....	40
5.6.	Redondeo Probabilístico.....	42
6.	CÁLCULO Y EVALUACIÓN DE LAS ESTIMACIONES.....	43



6.1. Estimación de la varianza compleja 43

6.2. Variables que identifican el diseño muestral complejo 44

6.3. Estadísticos asociados a los parámetros de interés 45

7. ANEXOS 47

 Anexo N°1: Categoría elegibilidad y códigos de disposición final (CDF) EBS 2025 47

 Anexo N°2: Sintaxis en R para la estimación de las variables de interés 48

8. REFERENCIAS..... 53



ÍNDICE DE TABLAS

Tabla 1: Resumen parámetros utilizados, tamaños y errores resultantes según escenario, EBS 2025	21
Tabla 2: Estimación de tamaños muestrales utilizando la muestra lograda de personas de 18 años o más en Casen 2024	22
Tabla 3: Tamaños y errores finales a nivel nacional, regional, urbano-rural, EBS 2025	24
Tabla 4: Tamaños y errores EBS 2023 y propuesta EBS 2025. Nacional Urbano - Rural	24
Tabla 5: Errores esperados asociados al parámetro de interés según tamaño muestral objetivo	25
Tabla 6: Distribución del marco de selección base y marco de selección filtrado EBS 2025	26
Tabla 7: Distribución del número de personas por región y área	28
Tabla 8. Número de personas según tipo de elegibilidad, EBS 2025	33
Tabla 9. Variables utilizadas en el modelo de no contacto EBS 2025	33
Tabla 10. Veintiles de ajuste por no contacto EBS 2025	35
Tabla 11. Variables utilizadas en el modelo de no respuesta EBS 2025	36
Tabla 12. Veintiles de ajuste por no respuesta EBS 2025	37
Tabla 13. Proyecciones de población al 30 de agosto 2025 según región y área	40
Tabla 14. Distribución de la muestra y factor de expansión calibrado por región y área, EBS 2025	41
Tabla 15. Errores muestrales por región y área para la variable Insatisfacción con la vida, EBS 2025	45
Tabla 16. Errores muestrales por región y área para la variable balance afectivo negativo, EBS 2023	46



ÍNDICE DE CUADROS

Cuadro 1. Parámetros utilizados para el cálculo de tamaño muestral actual17

Cuadro 2: Parámetros comunes utilizados en escenarios EBS 2025.....20



1. ANTECEDENTES

El diseño muestral de la EBS 2025 fue elaborado por el Instituto Nacional de Estadísticas (INE) en el marco del programa de trabajo conjunto acordado con el Ministerio de Desarrollo Social y Familia para esta versión de la encuesta. El presente informe es parte de los productos entregables de ese convenio.

El informe presenta el diseño muestral elaborado con el fin de responder a los objetivos propios de la encuesta, requerimientos que se deben considerar al momento de determinar un tamaño muestral adecuado. El tamaño muestral debe balancear en su diseño el presupuesto disponible y los errores de muestreo considerados como admisibles para fines del estudio. La conjugación de estos tres elementos — requerimientos, errores muestrales admisibles y presupuesto disponible— permite obtener tamaños muestrales óptimos.

La metodología del diseño muestral de EBS 2025 describe el conjunto de opciones metodológicas y de procedimientos estadísticos adoptados a lo largo del proceso de planificación, considerados los más apropiados para seleccionar una muestra probabilística y estadísticamente representativa de la población objetivo, que permita estimar los parámetros de interés en los dominios de estudio definidos para la encuesta.

Este informe contiene los siguientes capítulos:

- Características generales del diseño muestral
- Estimación de parámetros necesarios para el cálculo del tamaño muestral
- Estimación del tamaño muestral
- Selección de unidades muestrales
- Construcción de factores de expansión
- Cálculo y evaluación de las estimaciones



2. CARACTERÍSTICAS GENERALES DEL DISEÑO MUESTRAL

2.1. Objetivo

La EBS tiene por objetivo generar información sobre cómo las personas evalúan su propia vida respecto a diversas dimensiones, generando una batería de indicadores de bienestar que complementen la información entregada por la encuesta Casen 2024, así como también, actualizar el diagnóstico proporcionado por EBS 2023 sobre el bienestar y las oportunidades que tienen las personas en el país, además de obtener indicadores oficiales en diferentes dimensiones del bienestar.

2.2. Población objetivo

La población objetivo es el universo de personas de 18 años o más, cumplidos a la fecha de realización de Casen 2024, que residen en viviendas particulares en el territorio nacional.

2.3. Unidad de información

La EBS recoge información directamente de las personas de 18 años o más, lo que permite obtener un diagnóstico completo sobre su bienestar y las oportunidades que enfrentan, al mismo tiempo que, se obtiene información sobre la composición de sus hogares y de sus variables sociodemográficas.

2.4. Marco muestral

Las personas de 18 años o más entrevistadas en EBS 2025 son una muestra que proviene de los individuos pertenecientes a los hogares efectivamente entrevistados por la Encuesta Casen 2024. De esta manera, ambas encuestas comparten el mismo marco muestral basado en el Censo de Población y Viviendas de 2017 y actualizado al año 2021 (MMV 2021).

El marco muestral de viviendas utilizado en Casen 2024 fue elaborado a partir del MMV 2021 y está constituido por unidades primarias de muestreo (UPM) que corresponden a áreas geográficas homogéneas, en términos del número de viviendas particulares que las conforman, excluyendo viviendas de temporada y colectivas.

Las UPM presentan una estratificación geográfica, dada por la división político-administrativa, según la conformación región-provincia-comuna y desagregada, además según la división censal, clasificándolas en

áreas urbanas y rurales. A esta estratificación se suma la clasificación socioeconómica (con niveles bajo, medio y alto) construida a partir de un modelo que considera variables del Censo 2017³.

A través de las UPM se accede a las viviendas y hogares, cuyos integrantes son principalmente el objeto final de estudio. Para la selección de viviendas o unidades secundarias de muestreo (USM) - que son las unidades seleccionadas para la primera fase - se considera la actualización de las UPM seleccionadas en primera etapa mediante un trabajo de gabinete utilizando la información obtenida desde la Actualización Cartográfica Precensal 2023 y registros administrativos como, por ejemplo, Certificados de Recepción Final.

2.5. Dominios de estudio

La encuesta EBS 2025 se encuentra diseñada para obtener indicadores fiables con errores de muestreo aceptables definidos para los niveles nacional, nacional urbano, nacional rural y regional.

2.6. Objetivo de precisión y tamaño de muestra

Considerando un levantamiento telefónico con un tamaño objetivo ajustado de 13.035 personas, se espera obtener estimaciones del parámetro de interés con un error absoluto nacional de 1,0% y, un error relativo, de 8,9%. Con este tamaño los errores absolutos regionales no superan 3,7% y, los errores relativos, 43,8%. Pese a que existen errores relativos elevados, a partir del error estándar se observa que los valores a nivel regional cumplen con el estándar de calidad que utiliza el INE para la publicación de parámetros de interés.

2.7. Estrategia muestral

El diseño muestral de EBS 2025 implica la obtención de una muestra probabilística bifásica. La primera fase está dada por el diseño muestral de Casen 2024, que corresponde a una muestra probabilística, estratificada y bietápica, siendo los estratos conformados por la combinación comuna-área-nivel socioeconómico⁴. La segunda fase se compone de una muestra probabilística estratificada, siendo los estratos muestrales conformados por la combinación de región-área.

³ Documentos de trabajo: Estratificación socioeconómica del marco muestral de viviendas 2017. <https://www.ine.gob.cl/inicio/documentos-de-trabajo/documento/estratificación-socioeconómica-del-marco-muestral-de-viviendas-2017>

⁴ El nivel socioeconómico se incorpora para algunas combinaciones de comuna-área.

3. ESTIMACIÓN DEL TAMAÑO MUESTRAL

Debido a que en las diferentes regiones se pueden observar comportamientos distintos en los fenómenos de estudio y en la configuración de las unidades de primera y segunda etapa de selección, la estimación de parámetros de interés de la encuesta adquiere complejidad, así como también la estimación de sus estadísticos asociados, tales como: Errores muestrales, efectos de diseño, entre otros. Para representar estas características es necesario definir un tamaño muestral que capture la mayor parte de esta diversidad, de modo que, la encuesta refleje con la mayor precisión posible los aspectos sobre los que pretende dar respuesta.

En EBS 2025 el parámetro de interés que se busca estimar en la población objetivo es la tasa de insatisfacción de la vida, buscando que esta estimación se acerque lo más posible al parámetro poblacional⁵. La distancia entre el parámetro poblacional y su estimación⁶ puede ser usada como una medida que cuantifica cuán lejos o cerca se está del verdadero valor. Esto se puede expresar en términos probabilísticos y tomar los márgenes de error de la estimación, como el error absoluto y relativo, con cierto nivel de confianza.

A partir de lo anterior, en esta sección se presentan los elementos considerados para dar inicio al cálculo del tamaño muestral para EBS 2025 con el propósito de cumplir los objetivos propios de la encuesta, considerando factores como: variable de interés, el estimador del parámetro asociado a la variable, los niveles de estimación, los errores de muestreo tolerables y disponer de una fuente de información para obtener estimaciones que servirán para dar inicio al cálculo del tamaño muestral (levantamientos anteriores).

Además, es importante señalar una diferencia relevante respecto al trabajo realizado en EBS 2023, específicamente en relación con la estimación de la varianza y el efecto de diseño. Para esta instancia, la estimación del parámetro de interés considera que los datos provienen de un diseño muestral bifásico. Esto significa que se incorporan, en la estimación, tanto las variables de estratificación y conglomeración definidas en la encuesta Casen (primera fase) como los estratos definidos para EBS. De esta manera, la estimación de la varianza y del efecto de diseño reconoce que la muestra de segunda fase se extrajo condicional a un diseño complejo preexistente que corresponde a la primera fase.

Para esta versión de EBS, el MDSF ha definido previamente el tamaño muestral objetivo y con sobremuestreo. A partir de estos tamaños se realiza un proceso iterativo en el que se ajustan los errores muestrales y se

⁵ Las estimaciones puntuales buscan dar el mejor valor estimado individual de un parámetro, en comparación con la estimación de intervalo, que procede de especificar un rango de valores (OCDE, 2007).

⁶ Esta distancia o diferencia absoluta es denominada error de estimación, error de muestreo o margen de error.

evalúan diferentes tasas de no logro (TNL) para el cálculo del tamaño con sobremuestreo, hasta obtener un escenario que se adecúe a los requerimientos de la encuesta.

3.1. Etapas para la estimación del tamaño muestral para EBS 2025

El cálculo del tamaño muestral comienza con la obtención de parámetros de interés asociados a la última aplicación de la encuesta (EBS 2023) en los niveles de estimación definidos para el nuevo estudio (EBS 2025). Dado que la representatividad más desagregada de la encuesta es regional, se obtiene la estimación de la tasa de insatisfacción con la vida a ese nivel, junto con sus estadísticos asociados (error estándar y efecto del diseño).

El cálculo del tamaño muestral para EBS 2025 prevé una serie de procedimientos que corresponden, en principio, al paso desde un diseño bajo muestreo aleatorio simple a un diseño complejo, seguido de otros dos ajustes, determinados por la finitud de la población y la tasa de no logro, obteniendo así el tamaño con sobremuestreo.

En los párrafos siguientes se describe, en detalle, cada uno de los siete pasos realizados para definir el tamaño muestral con el total de personas a encuestar, en principio, a nivel de región, y luego a nivel de estrato de muestreo (región-área):

Paso 1: Estimación de parámetros

Se obtienen estimaciones, a nivel regional, de la tasa de insatisfacción con la vida junto con el error estándar y el efecto del diseño asociados a este indicador, usando los resultados del trabajo de campo de EBS 2023. A partir de estos parámetros se obtiene la cuasivarianza poblacional $S(r)_r^2$ según la ecuación (1):

$$S(r)_r^2 = \frac{m_r}{Def f(r)_r} \cdot SE(r)_r^2 \quad (1)$$

Donde:

- m_r : Número de personas logradas en la región r .
- $Def f(r)_r$: Efecto del diseño asociado a la tasa de insatisfacción con la vida (r) en la región r .
- $SE(r)_r$: Error estándar de la estimación de la tasa de insatisfacción con la vida (r) en la región r .

El efecto del diseño asociado al parámetro de interés se define como:

$$Def f(r) = \sigma^2(r)_{MC} / \sigma^2(r)_{MAS} \quad (2)$$

Donde:

- $\sigma^2(r)_{MC}$: Cálculo de la varianza bajo un diseño complejo.

$\sigma^2(r)_{MAS}$: Cálculo de la varianza bajo un muestreo aleatorio simple.

Es importante recordar que para EBS 2025, el efecto de diseño fue estimado considerando explícitamente que la varianza bajo diseño complejo corresponde a un diseño muestral bifásico. Este tipo de diseño implica que, en una primera fase, se seleccionan viviendas para la encuesta Casen —a partir de la selección de conglomerados en cada estrato—, y posteriormente, en una segunda fase, se seleccionan personas dentro de los hogares que respondieron en la primera fase, utilizando una estratificación propia de la encuesta.

En el cálculo del efecto de diseño (*Deff*) de una encuesta, es posible observar valores menores que 1 o mayores que lo esperado según el diseño muestral o lo observado en diferentes niveles de estimación. En las encuestas con diseños muestrales complejos, donde se estime un efecto de diseño menor que 1, puede deberse a que los algoritmos de cálculo de la varianza se basan en métodos aproximados como, por ejemplo, la linealización basada en una expansión de Taylor que se utilizó para aproximar la varianza de estimadores no lineales. Esta linealización puede generar una varianza menor que la obtenida bajo diseño aleatorio simple, resultando en un *Deff* menor que 1. Ante esta situación, y bajo el supuesto que una muestra polietápica tiene un mayor error de muestreo que una muestra bajo muestreo aleatorio simple, se justifica establecer un valor mínimo de 1 para el efecto de diseño utilizado en el cálculo de tamaños muestrales.

Por otra parte, en los casos en que quien investiga considere que los efectos de diseño estimados exceden de manera importante lo esperado, puede tomar la decisión de establecer cotas superiores para su uso en el cálculo de tamaños, bajo el supuesto que los altos valores observados se deben a factores particulares del diseño y de la muestra recolectada y, considerando además, que la inferencia se realiza sobre una submuestra particular de un gran número de muestras posibles obtenidas desde el levantamiento de la encuesta.

Paso 2: Cálculo de Tamaño bajo M.A.S.

Se obtienen los errores absolutos “objetivo” por región (d_{0r}) y luego, los errores relativos (e_{0r}) según una aproximación del cuantil de una distribución t de Student con nivel de confianza de 95% que, para efectos de la simulación, tiene un valor correspondiente a $t=2$. Luego, junto con la cuasivarianza calculada en la ecuación (1) se obtiene un tamaño muestral inicial a nivel regional (m_{0r}) considerando un muestreo aleatorio simple monoetápico.

$$m_{0r} = \frac{t_{1-\alpha/2}^2 \cdot S(r)_r^2}{d_{0r}^2} \quad (3)$$

Paso 3: Ajuste por efecto del diseño

Dado que el diseño muestral definido para EBS 2025 es un diseño complejo, el tamaño calculado bajo m.a.s. se debe ponderar por un valor que representa el efecto provocado en la varianza al muestrear bajo un diseño estratificado y bifásico en lugar de un m.a.s.

Así, el tamaño asociado al diseño complejo se obtiene desde la ecuación (4):

$$m_{1r} = m_{0r} \cdot Deff(r)_r \quad (4)$$

Donde:

- m_{0r} : Es el número de personas a encuestar en la región r bajo un muestreo aleatorio simple.
- m_{1r} : Es el número de personas a encuestar en la región r considerando un diseño complejo.
- $Deff(r)_r$: Es el efecto de diseño en la región r obtenido desde la estimación de la tasa de insatisfacción con la vida.

Generalmente, los valores de $Deff(r)_r$ son acotados al rango (1 – 5) basado en el supuesto que, por un lado, para un mismo tamaño muestral, una muestra polietápica tiene un mayor error de muestreo que bajo un muestreo aleatorio simple, lo que implica que $Deff$ no puede ser menor que 1 y, por otro, dado que la evidencia empírica en estudios de este tipo apunta a que la ganancia en precisión, producto de corregir el tamaño muestral por un efecto del diseño mayor a 5, es marginal respecto al incremento de los costos de levantamiento del nuevo tamaño muestral.

Paso 4: Corrección por población finita

El tamaño m_{1r} a su vez, es ajustado por un ponderador que da cuenta de la finitud de la población. Su forma de cálculo se observa en la ecuación (5):

$$m_{2r} = \frac{t_{1-\alpha/2}^2 \cdot S(r)_r^2 \cdot Deff(r)_r}{d_{0r}^2 + t_{1-\alpha/2}^2 \cdot Deff(r)_r \cdot S(r)_r^2 / M_r} = \frac{m_{1r}}{1 + m_{1r} / M_r} \quad (5)$$

Donde:

- m_{2r} : Es el total de personas a encuestar en la región r considerando un diseño complejo y asumiendo finitud de la población.
- M_r : Es el total de personas en la región r , según marco de selección.

Finalmente, este tamaño es el que determina la muestra objetivo regional, que se debe distribuir en cada uno de los estratos de muestreo definidos (ra).

Paso 5: Distribución por estrato de muestreo: región – área

Una vez determinado el tamaño muestral objetivo m_{2r} a nivel regional, se distribuye este valor, en forma proporcional al marco de selección, en cada uno de los estratos que las conforman.

Es importante mencionar que, el tamaño inicial calculado en cada estrato de muestreo puede ajustarse en la medida que estos no cumplan con tamaños mínimos definidos, por lo que, los tamaños regionales calculados en el paso 4 eventualmente podrían aumentar, caso en el cual, se deben recalcular los errores muestrales propuestos inicialmente.

Paso 6: Ajuste por no logro

Una vez definido el total de unidades a encuestar para estimar la tasa de insatisfacción con la vida, se debe tener en consideración que, en el trabajo de campo es posible no lograr algunas unidades por diversas razones, tales como: rechazos, imposibilidad de contactar al informante, entre otras. Por este motivo es necesario ajustar el tamaño muestral definido en el paso 5 por un ponderador que da cuenta de las posibles pérdidas de unidades muestrales.

Para calcular este ponderador, se utiliza información relacionada con el logro de la encuesta en levantamientos anteriores o de operaciones estadísticas de similar temática o método de recolección. El nivel de desagregación idóneo para el ponderador de ajuste es a nivel de estrato de muestreo. Obtenidos los tamaños con sobremuestreo a este nivel, los tamaños a nivel regional son calculados como la agregación de los tamaños de los estratos que las conforman.

Luego, el tamaño con sobremuestreo a nivel de estrato m_{3ra} es calculado en la ecuación (6):

$$m_{3ra} = \frac{m_{2ra}}{1 - tn l_{ra}} \quad (6)$$



Donde:

- m_{3ra} : Número total de personas a encuestar en el estrato de muestreo ra .
- tnl_{ra} : Tasa de no logro aplicada en el estrato de muestreo ra .

Paso 7: Estimación de la tasa de no respuesta a niveles agregados

Dado que el tamaño de muestra regional con sobremuestreo m_{3r} , es calculado como la suma o agregación de estratos de muestreo ra , del total de personas de cada una de las regiones, se puede despejar la tasa de no logro y estimar la tasa regional a partir de la ecuación (7):

$$tnl_r = \frac{m_{3r} - m_{2r}}{m_{3r}}$$

(7)

De forma análoga se estima la tasa de no logro a nivel nacional.

3.2. Parámetros utilizados para el cálculo del tamaño muestral

Para el cálculo del tamaño muestral se consideran los elementos presentados en el Cuadro 1:

Cuadro 1. Parámetros utilizados para el cálculo de tamaño muestral actual	
Parámetro	Descripción
Variable de diseño	$X_j = \begin{cases} 0 & , \text{ si la persona no está insatisfecha con su vida} \\ 1 & , \text{ si la persona está insatisfecha con su vida} \end{cases}$
Parámetro asociado	Tasa de insatisfacción con la vida (porcentaje de personas que respondieron “totalmente insatisfecho” e “insatisfecho” en la pregunta ¿Cuán satisfecho(a) o insatisfecho(a) está usted con su vida en este momento?
Estimador asociado	Estimador de razón de insatisfacción con la vida según resultados obtenidos de EBS 2023: $r = \frac{\text{Número de personas insatisfechas con su vida}}{\text{Número total de personas}}$
Tamaño de la población objetivo	1. Se disponen de proyecciones de población oficiales para el período de simulación de tamaño muestral de la encuesta. 2. Se dispone de la muestra lograda en Casen 2024 ⁷ de personas de 18 años o más por región y área.
Niveles de estimación	Nacional Nacional urbano, Nacional rural Regional

⁷ Debido a que a la fecha de la simulación del tamaño muestral y la selección de unidades no se encuentran publicados los resultados de Casen 2024, se utiliza una base preliminar para el cálculo de tamaño de la población objetivo.



Parámetro	Descripción
Errores de muestreo	Errores estándar, que cumplan con el umbral de aceptación definido por el estándar para la evaluación de la calidad de las estimaciones en encuestas de hogares (INE, 2020), definido por: $\text{Máximo ee tolerable} = \begin{cases} \sqrt[3]{(r^2)/9} & ; \quad r < 0,5 \\ \sqrt[3]{(1-r)^2/9} & ; \quad r \geq 0,5 \end{cases}$
Cuasivarianza	Obtenida desde el trabajo de campo de EBS 2023
Efecto del diseño	Obtenido desde el trabajo de campo de EBS 2023
Tasa de sobremuestreo	Tasas de no logro a nivel regional de EBS 2023 ajustadas.

Fuente: Instituto Nacional de Estadísticas.

3.3. Metodología de cálculo del tamaño muestral utilizando muestras complejas

La población objetivo, los niveles de estimación deseados y la cobertura geográfica de la encuesta EBS 2025, hacen necesaria la confección de un diseño eficiente para la selección de una muestra de personas de 18 años o más, en términos de ejecución presupuestaria e implementación operativa.

El muestreo aleatorio simple de personas necesitaría, en principio, de un marco de personas de 18 años o más con sus respectivas direcciones de domicilios, generado por un censo de población y viviendas que, por un lado, resultaría muy costoso de mantener actualizado y, por otro, las personas seleccionadas desde un marco de este tipo, se encontrarían tan dispersas geográficamente que, los costos asociados al levantamiento excederían la asignación presupuestaria, por tanto, el muestreo aleatorio simple iría en desmedro de la eficiencia deseada.

Por otra parte, en los diseños de muestra complejos, las unidades que finalmente componen la muestra (en este caso, personas de 18 años o más) se determinan después de varias etapas; en consecuencia, no es necesario mantener un marco de personas actualizado, lo cual reduce significativamente los costos. Este tipo de diseño implica una estructura de muestreo más sofisticada que, aunque eficiente operativamente, introduce efectos adicionales sobre la precisión de las estimaciones.

En este contexto, la metodología de cálculo del tamaño muestral utilizada en EBS 2025 se inicia con la determinación de un tamaño muestral bajo un muestreo aleatorio simple, fijando previamente el error absoluto a alcanzar con base en los resultados de encuestas anteriores. Posteriormente, se realizan tres ajustes secuenciales: el primero, que da cuenta del efecto del diseño; el segundo, que corresponde a un ajuste por población finita y; el tercero, referido al ajuste por la tasa de no logro.



Uno de los aspectos clave de estos ajustes es el efecto de diseño, definido como el cociente entre la estimación de la varianza bajo un diseño complejo versus la estimación que se obtendría con un diseño aleatorio simple. En términos prácticos, este indicador permite cuantificar cuánto debe aumentarse el tamaño muestral en un diseño complejo para alcanzar un nivel de precisión equivalente al de un muestreo aleatorio simple. Ahora bien, para determinar el tamaño de la muestra compleja se incorpora este factor de ajuste a partir del tamaño de una muestra aleatoria simple, lo que permite aproximarse al número de personas necesarias para que el diseño complejo proporcione la misma varianza (Kish, 1965).

De esta manera, utilizando los efectos que dan cuenta del diseño complejo, se calculan los tamaños necesarios que permitan obtener estimaciones con el grado de precisión calculado en el paso anterior. Para esta versión, los efectos de diseño utilizados están basados, en los obtenidos en el levantamiento de EBS 2023. Además, la corrección por finitud de la población es estimada a partir del marco de selección disponible para la realización de esta encuesta, razón por la que se utiliza la muestra lograda en la encuesta Casen 2024.

Posteriormente, luego de realizar un procedimiento mediante el cual se distribuye el tamaño de muestra de cada región al interior de los estratos de muestreo (región-área) donde se busca, por un lado, resguardar coherencia entre los tamaños muestrales y aquellos de la población (según el marco muestral utilizado), en términos de proporcionalidad y, por otro lado, que dicha distribución no difiriera mayormente de la distribución de las proyecciones de población al 15 de mayo de 2025 (correspondiente al período de simulación). Bajo esta distribución, los tamaños con sobremuestreo son calculados a partir de la tasa de no logro a nivel regional de EBS 2023.

Finalmente, para que la muestra sea eficiente, en términos de la cantidad de personas a seleccionar, la fracción de muestreo respecto del marco debe considerar el establecimiento de cotas, especialmente superiores a nivel regional. En este caso, las fracciones de muestreo no superan 27,7% a nivel regional.

3.4. Resumen proceso de simulación

En este apartado se presentan las diferentes simulaciones de distribución de tamaños muestrales realizadas para cumplir con los requerimientos establecidos por el MDSF: un tamaño muestral objetivo de 13.035 personas y un tamaño con sobremuestreo de 25.686 personas. La estructura del capítulo sigue la cronología de las simulaciones realizadas, junto con la metodología empleada en cada una de ellas, hasta llegar a la propuesta seleccionada.

Como punto de partida, se presentan los resultados del trabajo de campo de la EBS 2023 (escenario n°0), a partir del cual se obtiene la estimación de la tasa de insatisfacción con la vida, junto con sus estadísticos asociados: error estándar, coeficiente de variación y efecto de diseño, tanto a nivel nacional como regional.

Dado que el MDSF definió tanto el tamaño muestral objetivo (13.035 personas) como el tamaño muestral con sobremuestreo (25.686 personas) a nivel nacional, se elaboró solo una propuesta para el tamaño objetivo y tres alternativas para el tamaño con sobremuestreo.

Estas tres alternativas derivan del escenario n°1, el cual estima los tamaños muestrales objetivo utilizando la muestra lograda de personas de 18 años y más en Casen 2024 y, a partir de este escenario, se desarrollaron distintas propuestas que difieren principalmente en la tasa de no logro (TNL) utilizada, basada en los resultados de la EBS 2023.

- 1. En el escenario n°1.1, se propone utilizar la TNL a nivel regional de la EBS 2023.
- 2. En el escenario n°1.2, se plantea el uso de una TNL de 50% para cada región.
- 3. En el escenario n°1.3, se considera una TNL ajustada, con una reducción aproximada de 8% respecto a la tasa original a nivel regional, para aproximarse al tamaño con sobremuestreo definido por MDSF.

Finalmente, una vez definidos los tamaños muestrales con sobremuestreo para cada uno de los escenarios analizados, se evalúa cuál de estas alternativas se aproxima de mejor manera a la propuesta original del MDSF. Esta comparación permite seleccionar la opción más adecuada para cumplir con los requerimientos establecidos, asegurando así la coherencia entre los objetivos del estudio y la estrategia de muestreo planteada.

En el Cuadro 2 se presentan los parámetros comunes utilizados en los escenarios considerados.

Cuadro 2: Parámetros comunes utilizados en escenarios EBS 2025

Parámetro	Descripción
Parámetro a estimar	Tasa de insatisfacción con la vida
Variable de diseño	$X_j = \begin{cases} 0, & \text{si la persona no está insatisfecha con su vida} \\ 1, & \text{si la persona está insatisfecha con su vida} \end{cases}$
Estimador asociado	Estimador de razón: $r = \frac{\text{Número de personas insatisfechas con su vida}}{\text{Número total de personas}}$

Fuente: Instituto Nacional de Estadísticas.

La Tabla 1 presenta un resumen de los parámetros utilizados en cada una de las simulaciones, así como también, los tamaños y errores muestrales resultantes, por orden cronológico, comenzando con el escenario propuesto por el MDSF, para seguir con los resultados del trabajo de campo de EBS 2023 (escenario n°0) y luego, presentar el resto de los escenarios.



Tabla 1: Resumen parámetros utilizados, tamaños y errores resultantes según escenario, EBS 2025

Nro.	Niveles de estimación	Error Absoluto	Error Relativo	Efectos de Diseño	Tamaño Muestral Objetivo	Tamaño Muestral Objetivo (ajustado) ⁸	Tasa de no logro	Tamaño con sobremuestreo
1.1	Nacional	0,9%	8,5%	2.9	12.943	13.035	TNL	30.868
	Regional	1,8% - 3,4%	14,4% - 41,4%	1,6-2,1			EBS 2023	
1.2	Nacional	0,9%	8,5%	2.9	12.943	13.035	TNL	25.929
	Regional	1,8% - 3,4%	14,4% - 41,4%	1,6-2,1			ajustada a 50%	
1.3	Nacional	0,9%	8,5%	2.9	12.943	13.035	TNL con reducción de ~8%	25.686
	Regional	1,8% - 3,4%	14,4% - 41,4%	1,6-2,1				

Fuente: Instituto Nacional de Estadísticas.

⁸ Los errores absolutos y relativos son calculados con el tamaño objetivo ajustado.



3.5. Escenario escogido

3.5.1. Escenario 1.3

Tabla 2: Estimación de tamaños muestrales utilizando la muestra lograda de personas de 18 años o más en Casen 2024

Región	Personas ≥18 años Casen 2024	Proyección Personas ≥18 años 15/05/25	Tasa insatis- facción con la vida	Error absoluto propuesto	Error relativo propuesto	Tamaño bajo muestreo aleatorio simple	Efecto Diseño	Tamaño ajustado por efecto de diseño	Tamaño ajustado por finitud	Tamaño Objetivo Ajustado	Error Absoluto ajustado	Error Relativo ajustado	Tamaño con Sobre- muestreo
Nivel País	157.028	15.740.878	11,0%	0,9%	8,5%	8.021	2,9	14.175	12.943	13.035	1,0%	8,9%	25.686
Urbano	128.763	13.942.394	11,2%	1,0%	9,1%	7.103	2,8	12.553	11.462	10.998	1,1%	9,7%	21.753
Rural	28.265	1.798.484	10,1%	2,3%	23,1%	918	3,0	1.622	1.481	2.037	2,1%	20,6%	3.933
Arica y Parinacota	5.740	201.114	11,3%	3,4%	30,3%	349	1,8	628	567	580	3,6%	31,5%	1.207
Tarapacá	6.672	303.473	10,2%	3,1%	30,4%	379	1,8	682	619	644	3,2%	31,2%	1.331
Antofagasta	7.268	549.387	10,9%	3,3%	30,3%	361	2,1	758	687	721	3,4%	31,1%	1.494
Atacama	6.649	239.897	8,8%	3,0%	34,0%	359	1,8	646	589	596	3,1%	35,4%	1.186
Coquimbo	7.024	682.023	7,8%	2,4%	30,4%	515	1,7	876	779	779	2,5%	32,2%	1.613
Valparaíso	17.299	1.608.376	10,5%	2,4%	22,9%	634	1,7	1.078	1.015	1.015	2,5%	23,6%	1.930
Metropolitana	28.796	6.636.178	12,2%	1,8%	14,4%	1.385	1,6	2.216	2.058	2.058	1,8%	14,9%	4.209
O'Higgins	10.794	802.272	10,2%	2,6%	25,6%	532	1,7	904	835	835	2,7%	26,6%	1.574
Maule	10.300	916.032	10,9%	2,6%	24,1%	552	1,7	938	860	860	2,7%	25,2%	1.544
Ñuble	7.272	412.143	11,4%	3,2%	28,3%	387	1,9	735	668	668	3,4%	29,7%	1.252
Biobío	16.404	1.321.876	10,8%	2,5%	23,6%	592	1,7	1.006	948	948	2,6%	24,3%	1.934
La Araucanía	10.381	799.160	10,4%	2,6%	25,4%	530	1,7	901	830	830	2,8%	26,5%	1.547
Los Ríos	6.530	324.619	10,7%	3,0%	27,8%	427	1,7	726	654	654	3,1%	29,3%	1.253
Los Lagos	8.343	715.405	8,8%	2,7%	31,2%	425	2,1	893	807	807	2,9%	32,8%	1.565
Aysén	3.169	82.580	8,5%	3,4%	40,4%	270	2,0	540	462	462	3,7%	43,6%	877
Magallanes	4.387	146.343	6,8%	2,8%	41,4%	324	2,0	648	565	578	3,0%	43,8%	1.170

Fuente: Instituto Nacional de Estadísticas.

En la Tabla 2 se presenta la distribución de tamaños muestrales para el escenario n°1.3 que corresponde al escenario elegido.

Este escenario se basa en la estimación de parámetros obtenidos desde el trabajo de campo de EBS 2023 (escenario n°0). El error absoluto propuesto fue ajustado para cumplir el estándar de calidad y con este error se obtuvo un tamaño bajo muestreo aleatorio simple de 8.021 personas a nivel nacional.

El efecto de diseño utilizado corresponde al resultante en EBS 2023. A partir de este valor, se obtiene un tamaño objetivo ajustado de 14.175 personas. Posteriormente, se aplica el ajuste por finitud, que considera el tamaño del marco de selección construido a partir de la muestra lograda en Casen 2024, y se obtiene un tamaño de 12.943 personas a nivel nacional.

Cuando se ajusta la distribución por estrato muestral (región-área) en forma proporcional a la estimación de personas según Casen 2024, uno de los primeros criterios a considerar es contar con, al menos sesenta (60) observaciones en el estrato, con la finalidad de posibilitar la obtención de estimaciones fiables según lo establecido por el estándar para la evaluación de la calidad de las estimaciones del INE. A partir de lo anterior, se obtiene el tamaño objetivo resultante de 13.035 personas, con un error absoluto de 1,0% y un error relativo de 8,9%, a nivel nacional. Si bien, a nivel regional se observan errores relativos elevados, a partir del error estándar se observa que estos valores cumplen con el estándar de calidad que utiliza el INE para la publicación de estadísticas precisas y confiables.

Luego, para obtener el tamaño con sobremuestreo, a los tamaños objetivos resultantes por estrato se aplican diferentes tasas de no logro (TNL), con el fin de acercarse lo más posible al valor entregado por el MDSF que corresponde a 25.684 personas. Las tasas de no logro aplicadas corresponden a una modificación de las tasas de no logro (TNL) obtenidas en la EBS 2023, desagregadas a nivel regional. En concreto, se aplicó una reducción de 8% a las TNL en la mayoría de las regiones, mientras que, en las regiones de Antofagasta y Metropolitana la disminución fue de 9%, y en Tarapacá y Coquimbo, de 10%. A partir de estas tasas ajustadas, se obtuvo un tamaño muestral con sobremuestreo de 25.686 personas.

3.6. Tamaño muestral propuesto

El tamaño calculado a nivel regional se debe distribuir por área (urbana – rural). Para esto, se cuenta con la distribución resultante de Casen 2024 o bien, con las proyecciones de población estimadas al 15 de mayo de 2025 (fecha que coincide con el periodo simulaciones de tamaño de EBS 2025) al mismo nivel.

El tamaño objetivo ajustado del escenario elegido (n°1.3) definido para obtener estimaciones de la tasa de insatisfacción con la vida con suficiente grado de precisión a nivel regional, alcanza 13.035 personas a nivel nacional, cuyos errores absolutos y relativos asociados son de 1,0% y 8,9%, respectivamente. Aplicando las

tasas de no logro de EBS 2023 a nivel regional con los ajustes explicados previamente se obtiene el tamaño con sobremuestreo que alcanza 25.686 personas.

La Tabla 3 muestra los tamaños objetivo ajustados junto con sus errores muestrales asociados y el tamaño con sobremuestreo a nivel nacional y regional, desagregados por área urbano-rural.

Tabla 3: Tamaños y errores finales a nivel nacional, regional, urbano-rural, EBS 2025

Región	Error Absoluto	Error Relativo	Tamaño objetivo			Tamaño con sobremuestreo		
			Total	Urbano	Rural	Total	Urbano	Rural
Nivel País	1,0%	8,9%	13.035	10.998	2.037	25.686	21.753	3.933
Arica y Parinacota	1,1%	9,7%	580	520	60	1.207	1.082	125
Tarapacá	2,1%	20,6%	644	584	60	1.331	1.207	124
Antofagasta	3,6%	31,5%	721	661	60	1.494	1.370	124
Atacama	3,2%	31,2%	596	536	60	1.186	1.067	119
Coquimbo	3,4%	31,1%	779	637	142	1.613	1.319	294
Valparaíso	3,1%	35,4%	1.015	932	83	1.930	1.772	158
Metropolitana	2,5%	32,2%	2.058	1.984	74	4.209	4.058	151
O'Higgins	2,5%	23,6%	835	627	208	1.574	1.182	392
Maule	1,8%	14,9%	860	639	221	1.544	1.147	397
Ñuble	2,7%	26,6%	668	472	196	1.252	885	367
Biobío	2,7%	25,2%	948	844	104	1.934	1.722	212
La Araucanía	3,4%	29,7%	830	597	233	1.547	1.113	434
Los Ríos	2,6%	24,3%	654	477	177	1.253	914	339
Los Lagos	2,8%	26,5%	807	598	209	1.565	1.160	405
Aysén	3,1%	29,3%	462	372	90	877	706	171
Magallanes	2,9%	32,8%	578	518	60	1.170	1.049	121

Fuente: Instituto Nacional de Estadísticas

Luego, a partir de las desagregaciones de tamaños por estrato geográfico (región-área) y de los resultados del trabajo de campo de EBS 2023 se obtienen, por agregación, tamaños asociados para los niveles Nacional Urbano y Nacional Rural y, a partir de estos, se calculan los errores absolutos y relativos resultantes para dichos niveles que se muestran en la Tabla 4.

Tabla 4: Tamaños y errores EBS 2023 y propuesta EBS 2025. Nacional Urbano - Rural

	Resultados EBS 2023				Propuesta EBS 2025			
	Muestra efectiva	Tasa Insatisfacción con la vida	Error Absoluto	Error Relativo	Tamaño Objetivo	Error Absoluto	Error Relativo	Tamaño con Sobremuestreo
Nivel País	11.234	11,0%	1,0%	9,2%	13.035	1,0%	8,9%	25.686
Urbano	9.360	11,2%	1,1%	9,8%	10.998	1,1%	9,7%	21.753



Rural	1.874	10,1%	2,4%	24,1%	2.037	2,1%	20,6%	3.933
-------	-------	-------	------	-------	-------	------	-------	-------

Fuente: Instituto Nacional de Estadísticas.

A nivel nacional urbano, con un tamaño objetivo ajustado de 10.998 personas se obtiene una estimación para la tasa de insatisfacción con la vida con un error absoluto de 1,1% y un error relativo de 9,7%.

A nivel nacional rural, con un tamaño objetivo de 2.037 personas se obtiene una estimación para la tasa de insatisfacción con la vida con un error absoluto de 2,1% y un error relativo de 20,6%.

3.7. Errores esperados

La Tabla 5 presenta los errores esperados, asociados a la estimación del parámetro de interés para un tamaño muestral objetivo de 13.035 personas⁹.

Tabla 5: Errores esperados asociados al parámetro de interés según tamaño muestral objetivo

Región	Tasa insatisfacción con la vida EBS 2023	Error absoluto ajustado EBS 2025	Error relativo ajustado EBS 2025	Tamaño ajustado por finitud EBS 2025	Tamaño objetivo ajustado EBS 2025	Tamaño con sobremuestreo EBS 2025
Nacional	11,0%	1,0%	8,0%	12.943	13.035	25.686
Arica y Parinacota	11,3%	3,6%	31,5%	567	580	1.207
Tarapacá	10,2%	3,2%	31,2%	619	644	1.331
Antofagasta	10,9%	3,4%	31,1%	687	721	1.494
Atacama	8,8%	3,1%	35,4%	589	596	1.186
Coquimbo	7,8%	2,5%	32,2%	779	779	1.613
Valparaíso	10,5%	2,5%	23,6%	1.015	1.015	1.930
Metropolitana	12,2%	1,8%	14,9%	2.058	2.058	4.209
O'Higgins	10,2%	2,7%	26,6%	835	835	1.574
Maule	10,9%	2,7%	25,2%	860	860	1.544
Ñuble	11,4%	3,4%	29,7%	668	668	1.252
Biobío	10,8%	2,6%	24,3%	948	948	1.934
La Araucanía	10,4%	2,8%	26,5%	830	830	1.547
Los Ríos	10,7%	3,1%	29,3%	654	654	1.253
Los Lagos	8,8%	2,9%	32,8%	807	807	1.565
Aysén	8,5%	3,7%	43,6%	462	462	877
Magallanes	6,8%	3,0%	43,8%	565	578	1.170

Fuente: Instituto Nacional de Estadísticas.

⁹ La estimación de los errores se obtiene a partir del tamaño objetivo y no del tamaño con sobremuestreo, pues dadas las pérdidas naturales de unidades muestrales, se espera obtener un tamaño de 13.035 personas.

4. SELECCIÓN DE UNIDADES MUESTRALES

La selección de unidades muestrales se realiza en forma sistemática y con igual probabilidad al interior de cada estrato de muestreo (región - área). La selección se realiza desde el marco muestral actualizado a mayo de 2025, conformado por personas de 18 años o más logradas en Casen 2024, excluyendo a las personas que pertenecen al servicio doméstico puertas adentro.

Para la construcción del marco de selección, además de la exclusión de personas pertenecientes al servicio doméstico, fueron excluidas aquellas unidades que no presentaban información de contacto válida, es decir, personas sin nombre y/o número de teléfono de contacto válido. Debido a este filtro aplicado en la construcción del marco de selección, el listado de personas se redujo de 171.373 a 157.022 unidades.

En la tabla 6 es posible observar que el porcentaje de personas sin información de contacto válida alcanza 8,4% del total de personas de 18 años o más en Casen 2024. A su vez, la distribución de este fenómeno no es homogéneo entre regiones por lo que, en la construcción de los factores de expansión, esta situación será abordada a través de un ajuste en los ponderadores.

Tabla 6: Distribución del marco de selección base y marco de selección filtrado EBS 2025.

Región	Marco de selección completo	Personas sin nombre y/o teléfono válido	Marco de selección filtrado ¹⁰	Porcentaje del marco sin información de contacto
Nacional	171.373	14.351	157.022	8,4%
Arica y Parinacota	6.487	747	5.740	11,5%
Tarapacá	7.124	452	6.672	6,3%
Antofagasta	8.072	807	7.265	10,0%
Atacama	7.067	418	6.649	5,9%
Coquimbo	8.073	1.049	7.024	13,0%
Valparaíso	18.376	1.077	17.299	5,9%
Metropolitana	31.860	3.064	28.796	9,6%
O'Higgins	11.520	726	10.794	6,3%
Maule	11.640	1.342	10.298	11,5%
Ñuble	7.558	286	7.272	3,8%
Biobío	17.840	1.436	16.404	8,0%
La Araucanía	11.300	919	10.381	8,1%
Los Ríos	7.276	746	6.530	10,3%
Los Lagos	9.015	673	8.342	7,5%
Aysén	3.331	162	3.169	4,9%
Magallanes	4.834	447	4.387	9,2%

¹⁰ El total de personas disponibles en el marco de selección se redujo respecto al total utilizado en las simulaciones de tamaño muestral debido a que, tras las validaciones de la base Casen 2024, hubo 6 personas excluidas de la base de datos.



Fuente: Instituto Nacional de Estadísticas

Respecto a la selección de personas, esta etapa es implementada en el software R, utilizando el paquete "Sampling", fijando una semilla aleatoria, a fin de que pueda ser replicable en cualquier momento.

Sean M_{ra} el número de personas encontradas en el estrato y m_{ra}^{sel} el número de personas a seleccionar. Para la selección de m_{ra}^{sel} personas, el software ejecuta los siguientes pasos:

Paso 1

En primera instancia se ordenan geográficamente y en forma ascendente todas las personas según región, comuna, área, identificador de la UPM, sexo y edad¹¹.

Paso 2

Al interior de cada estrato de muestreo se calcula el período (k) que corresponde a:

$$k = M_{ra}/m_{ra}^{sel}$$

Notar que " k " puede ser un número real, no entero (puede tener decimales).

Paso 3

Luego se determina el arranque " A " o primera selección, que corresponde a una semilla aleatoria propia para la encuesta.

Paso 4

Posteriormente se suma sucesivamente el período " k " al arranque " A " para obtener distintos valores, los que dan origen a la selección de unidades de la siguiente forma: " A ", " $A + k$ ", " $A + 2k$ ", " $A + 3k$ ", ..., " $A + (m_{ra}^{sel} - 1)k$ ".

La primera persona seleccionada es " A " y es un número entero, la segunda es el redondeo de " $A + k$ ", la tercera es el redondeo de " $A + 2k$ " y así sucesivamente, hasta la m_{ra}^{sel} selección, dada por el redondeo de " $A + (m_{ra}^{sel} - 1)k$ ".

¹¹ En la sintaxis enviada a MDSF la variable "correlativo" replica este ordenamiento.



Luego, la probabilidad de inclusión de la j – ésima persona dentro del estrato em está dada por:

$$P_{re}(j) = \frac{m_{ra}^{sel}}{M_{ra}} \tag{1}$$

Donde:

- m_{ra}^{sel}
- : Número de personas seleccionadas en el estrato ra
- M_{ra}
- : Número actualizado de personas en el marco de selección en el estrato ra .

La Tabla 7 presenta la distribución del total de personas seleccionadas, según región y área (Urbano - Rural) obtenida a partir de la distribución del marco de personas de 18 años o más logradas en Casen 2024.

Tabla 7: Distribución del número de personas por región y área

Región	Área	Total personas	
		Marco Casen 2024	Selección EBS 2025
Total País		157.022	25.680 ¹²
Arica y Parinacota	Urbano	5.374	1.082
	Rural	366	125
Tarapacá	Urbano	6.298	1.207
	Rural	374	124
Antofagasta	Urbano	6.837	1.367
	Rural	428	124
Atacama	Urbano	6.071	1.067
	Rural	578	119
Coquimbo	Urbano	5.399	1.319
	Rural	1.625	294
Valparaíso	Urbano	14.559	1.772
	Rural	2.740	158
Metropolitana	Urbano	26.930	4.058
	Rural	1.866	151
O'Higgins	Urbano	7.694	1.182
	Rural	3.100	392
Maule	Urbano	7.113	1.146
	Rural	3.185	396
Ñuble	Urbano	5.082	885
	Rural	2.190	367
Biobío	Urbano	13.520	1.722
	Rural	2.884	212
La Araucanía	Urbano	7.092	1.113

¹² El total de personas seleccionadas se redujo debido a que tras las validaciones de la base Casen 2024 hubo 6 personas que fueron excluidas de la base de datos.



Región	Área	Total personas	
		Marco Casen 2024	Selección EBS 2025
Los Ríos	Rural	3.289	434
	Urbano	4.523	914
	Rural	2.007	339
Los Lagos	Urbano	5.925	1.160
	Rural	2.417	404
Aysén	Urbano	2.348	706
	Rural	821	171
Magallanes	Urbano	3.994	1.049
	Rural	393	121

Fuente: Instituto Nacional de Estadísticas.

5. CONSTRUCCIÓN FACTORES DE EXPANSIÓN

Los análisis y estudios basados en encuestas de hogares con diseño muestral probabilístico utilizan un ponderador en la estimación de parámetros de interés para que estos tengan validez sobre la población objetivo. Este ponderador guarda relación con las probabilidades de selección de las distintas unidades de muestreo y da cuenta del número de personas de la población que representa cada individuo que participa en la encuesta. Este ponderador es conocido como factor de expansión.

Para EBS 2025 se desarrollan factores que expanden a la población regional y se calculan para cada persona presente en la muestra. El cálculo de estos factores responde al diseño muestral bifásico en cuya primera fase se obtiene un factor asociado a las viviendas logradas en Casen 2024 y, en la segunda fase, se calcula un factor que da cuenta de la selección de personas, realizada en forma sistemática y con igual probabilidad al interior de cada estrato de muestreo (región-área).

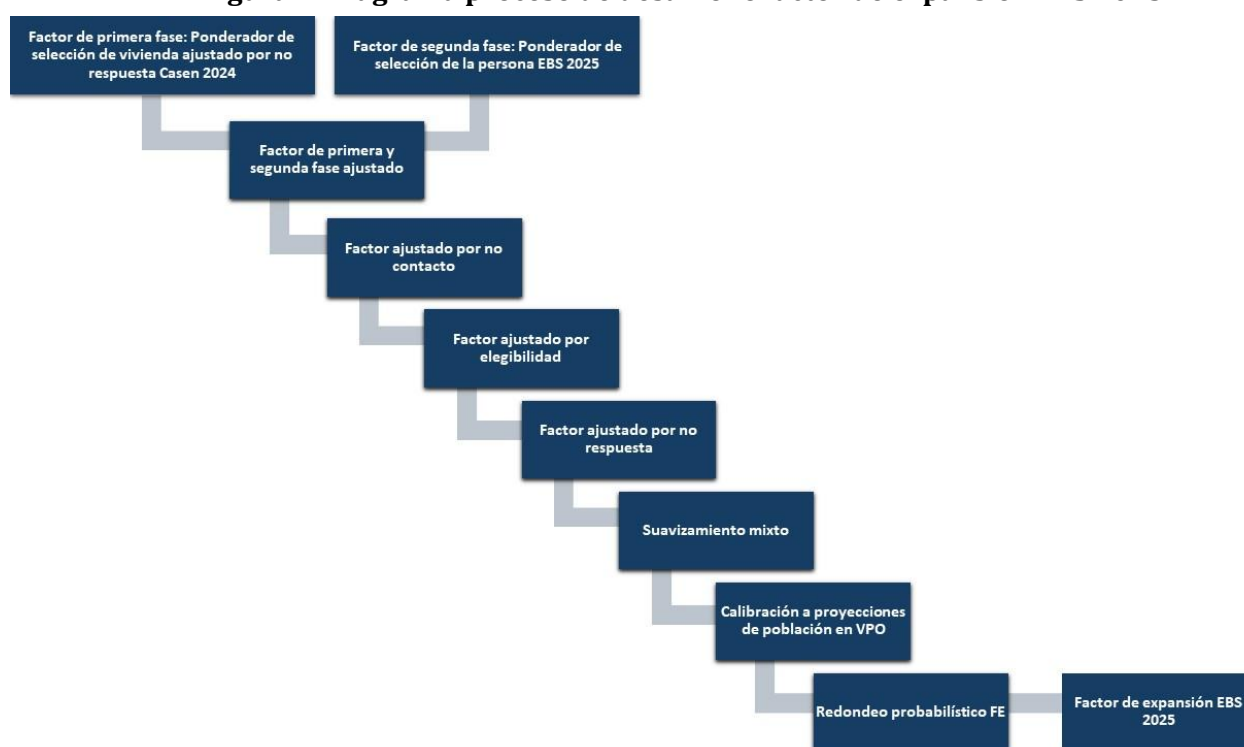
La metodología de cálculo de los factores de expansión consiste en la obtención secuencial de una serie de ponderadores que, como se menciona en párrafo precedente, se inicia con el factor de expansión de primera fase, que corresponde al ponderador de no respuesta de las viviendas logradas en Casen 2024 y que da cuenta de su diseño muestral. Luego, se obtiene el factor de expansión de segunda fase el que refleja cuántas personas en el marco de selección representa la persona seleccionada en la muestra de EBS.

A continuación, se obtiene el producto entre el factor de primera y segunda fase y se aplica una serie de ajustes según el resultado del trabajo de campo (no contacto y no elegibilidad) y la definición del logro de las personas (no respuesta), este último con el objetivo de reducir el sesgo asociado a la falla en conseguir la cooperación de las personas seleccionadas en la muestra. Al factor (o ponderador) resultante se le aplica, un suavizamiento, según la variabilidad del factor y finalmente, una calibración, con el fin de cuadrar las estimaciones poblacionales obtenidas de la encuesta con las proyecciones de población desarrolladas por el INE y que son utilizadas en todas sus encuestas de hogares.

Para la calibración se utiliza la técnica *Raking* que permite ajustar simultáneamente por diversas variables manteniendo el diseño muestral. Esta técnica reduce el sesgo que se produce por errores no muestrales, provocados por la no respuesta y/o la imperfección de los marcos muestrales y, además, mejora la precisión de las estimaciones. En el proceso de calibración, se utiliza la información auxiliar proveniente de las proyecciones de población de 18 años y más según sexo y edad para distintos niveles de desagregación, considerando solo la población residente en viviendas particulares.

La Figura 1 muestra el flujo del proceso del cálculo de los factores de expansión para EBS 2025 y a lo largo del capítulo se detalla cada una de las etapas del proceso de construcción de estos.

Figura 1. Diagrama proceso de desarrollo factor de expansión EBS 2025



Fuente: Instituto Nacional de Estadísticas.

5.1. Factor de expansión de primera fase

En Casen 2024 las unidades muestrales se obtienen desde el MMV 2021 a partir de un diseño probabilístico, estratificado y bietápico, donde los estratos muestrales están conformados por la combinación comuna – área – NSE, donde aplique, y al interior de cada estrato, los conglomerados, son seleccionados en forma sistemática con probabilidad proporcional al tamaño, según el número de viviendas que contienen. Por su

parte, las viviendas o unidades secundarias de muestreo son seleccionadas en forma aleatoria y con igual probabilidad al interior de cada UPM elegida.

Para el cálculo del factor de expansión de primera fase de EBS se considera el diseño muestral de Casen 2024 donde, al factor de expansión, se le aplican ajustes por elegibilidad y no respuesta para cada vivienda lograda en la muestra. De esta forma, el factor de expansión de primera fase ($F1_{hij}$) corresponde al ponderador ajustado por no respuesta de Casen 2024 (w'_{hij}^{NR}) que es obtenido a partir del cálculo de una razón de ajuste que permite que los ponderadores asociados a las viviendas elegibles que no respondieron se redistribuyan entre las viviendas elegibles que sí respondieron.

Por lo tanto, el factor de primera fase de EBS corresponde al factor ajustado por no respuesta para las viviendas logradas en Casen 2024 y se expresa según la ecuación (9):

$$F_{hij} = w'_{hij}^{NR} = \hat{R}_{g,R} \cdot w'_{hij} \quad (9)$$

Donde,

hij : Índice del estrato de muestreo, conglomerado y vivienda de Casen 2024

w'_{hij}^{NR} : Ponderador de selección de viviendas de Casen 2024 ajustado por no respuesta.

5.2. Factor de expansión de segunda fase

Dado que EBS se trata de una encuesta bifásica, luego de obtener los ponderadores asociados a las viviendas desde Casen 2024 es necesario calcular los ponderadores de las personas seleccionadas en la muestra.

5.2.1. Ponderador de selección de personas como inverso de la probabilidad de selección

Este ponderador corresponde al inverso de la probabilidad de selección de la persona para la muestra de EBS según el estrato de muestreo (región – área). Se interpreta como el número de personas, en la población de personas de 18 años y más, que representa la persona seleccionada en el marco de selección obtenido desde Casen 2024.

El ponderador de selección de personas es calculado como el inverso de la probabilidad de selección de la persona k perteneciente al estrato ra . Esto es:

$$F2_{ra,k} = \frac{M_{ra}}{m_{ra\ sel}} \quad (10)$$

Donde,

ra,k : Índice del estrato de muestreo región-área y de la persona seleccionada, respectivamente.

- $F2_{ra,k}$: Ponderador de selección de la persona k en el estrato ra .
- M_{ra} : Total de personas disponibles para seleccionar en el estrato ra , según el levantamiento de Casen 2024.
- $m_{ra\ sel}$: Total de personas seleccionadas en EBS 2025 en el estrato ra .

Sin embargo, el marco de selección no contempla todas las personas de 18 años o más que fueron levantadas en Casen 2024 debido a que no todos cumplían con la condición de contar con información de contacto válida. Esto implica que 8,4% de las personas fueron excluidas del marco de selección lo que hace necesario realizar un ajuste que dé cuenta de esta omisión con la finalidad de conservar el stock original de personas de 18 años y más de Casen 2024. Para ello, a nivel de estrato, se estima una razón entre la cantidad de personas seleccionables desde Casen 2024 y la cantidad total de personas de 18 años que respondieron dicha encuesta.

$$R_{ra} = \frac{M_{ra\ Casen}}{M_{ra}} \quad (11)$$

Donde,

- ra : Índice del estrato de muestreo región-área
- $M_{ra\ Casen}$: Total de personas en el estrato ra según el levantamiento de Casen 2024.

De este modo, el ponderador de selección de segunda fase está dado por:

$$F2_{ra,k\ aj} = F2_{ra,k} * R_{ra} \quad (12)$$

5.2.2. Producto del factor de primera y segunda fase

Este ajuste da cuenta de la selección de personas considerando ambas fases de muestreo, es decir, contempla la selección de la k -ésima persona para EBS, dado que fue entrevistada y lograda en Casen 2024. Este factor corresponde al factor teórico y se expresa según la ecuación (13):

$$w_{ra,k} = F1_{hij} * F2_{ra,k\ aj} \quad (13)$$

Donde,

- $w_{ra,k}$: Factor de selección de personas considerando la primera y segunda fase.

5.3. Ajustes por no contacto, elegibilidad y no respuesta

En principio la muestra incluye personas elegibles ya que estas pertenecen a hogares encuestados en Casen 2024. Sin embargo, durante el levantamiento hay situaciones en las que no es posible establecer contacto con la persona seleccionada o bien, hay personas que dejan de ser consideradas elegibles por diferentes motivos como, por ejemplo, dejar de pertenecer al hogar encuestado en Casen.

Estas situaciones se identifican en el cuestionario hoja de ruta a través de diferentes códigos de disposición final (CDF) donde, en el caso de EBS 2025, al tratarse de un levantamiento telefónico, las personas seleccionadas son clasificadas en elegibles y no elegibles¹³. La tabla 8 muestra la distribución de la elegibilidad en EBS 2025.

Tabla 8. Número de personas según tipo de elegibilidad, EBS 2025

Elegibilidad	Total de personas	Porcentaje del total
Total	25.680	
Elegibles entrevistadas	13.600	53,0%
Elegibles no entrevistadas ¹⁴	5.388	21,0%
Elegibles no contactadas	5.145	20,0%
No elegibles	1.547	6,0%

Fuente: Instituto Nacional de Estadísticas.

5.3.1. Ajuste por no contacto

Luego del trabajo de campo de EBS 2025, las personas no contactadas corresponden a 20% de la muestra gestionada. Entre las principales razones del no contacto se encuentran: el número telefónico no existe; el teléfono marca, pero no hay respuesta o el teléfono se encuentra apagado o fuera del área de cobertura. Si bien las personas no contactadas entran en la categoría de elegibles, no se sabe con certeza si efectivamente pertenecen a este grupo, es por este motivo que es necesario aplicar un ajuste que distribuye los ponderadores asociados a las personas no contactadas entre aquellas que sí lo fueron.

Para ello se estima la probabilidad de que la persona sea contactada mediante una regresión logística aplicada a las 25.680 personas seleccionadas utilizando variables de Casen 2024 y la hoja de ruta de EBS 2025. Luego, estas probabilidades de contacto son ordenadas de menor a mayor y las personas son agrupadas en veintiles de ajuste a nivel nacional.

En la tabla 9 se presentan las variables utilizadas en el modelo de no contacto.

Tabla 9. Variables utilizadas en el modelo de no contacto EBS 2025

Variable	Fuente
Horario de contacto (mañana, tarde o noche)	Hoja de ruta EBS 2025
Sexo persona recolectora	Hoja de ruta EBS 2025
Proporción de encuestas logradas sobre el total de encuestas asignadas a la persona recolectora	Hoja de ruta EBS 2025
Número de llamados (logaritmo)	Hoja de ruta EBS 2025

¹³ En la sección de Anexos se presenta tabla con el detalle de los Códigos de disposición final y su categoría de elegibilidad.

¹⁴ Esta categoría incluye a las personas con entrevista parcial, rechazo y otros motivos de no entrevista.



Variable	Fuente
Región	Casen 2024
Persona informante	Casen 2024
Persona ocupada	Casen 2024
Dificultad para hacer o recibir llamadas	Casen 2024
Condición permanente o de larga duración: Dificultad mental o intelectual	Casen 2024
Participación en alguna organización social, política o corporativa	Casen 2024
Pobreza por ingresos	Casen 2024
Quintil autónomo nacional	Casen 2024
Presencia de mayores de 60 años en el hogar	Casen 2024
Hogar carente en seguridad	Casen 2024
Tipo de vivienda	Casen 2024
Número de personas en el hogar (excluye servicio doméstico puertas adentro)	Casen 2024
Vivienda a menos de 8 cuadras o 1 km del transporte público	Casen 2024
Condición de tenencia de la vivienda	Casen 2024
Índice de saneamiento en la vivienda (acceso a servicios básicos)	Casen 2024
Acceso a internet	Casen 2024
Tipo de hogar	Casen 2024

Fuente: Instituto Nacional de Estadísticas.

Los veintiles de ajuste tienen como característica que agrupan personas de similares características por lo que, las personas contactadas podrán representar a aquellas que no lo fueron y, por lo tanto, el peso de las 5.145 personas no contactadas es distribuido entre aquellas que sí lo fueron, a través de una razón de ajuste realizada al interior de cada veintil utilizando el factor de selección de primera y segunda fase como se muestra en la ecuación (14).

En el numerador se incluye la estimación del total de personas en la muestra, y en el denominador la estimación del total de personas que fueron contactadas.

$$\hat{R}_{g,c} = \frac{\sum_{i \in \Phi_g} \sum_{j \in \theta_g} w_{ra,k}}{\sum_{i \in \Phi_g} \sum_{j \in \theta_{g,c}} w_{ra,k}} \quad (14)$$

Donde,

- g : : Índice del grupo para la corrección de no contacto (veintil).
- Φ_g : : Conjunto de personas asignadas al veintil g .
- θ_g : : Conjunto de personas pertenecientes al veintil g en la muestra.
- $\theta_{g,c}$: : Conjunto de personas pertenecientes al veintil g y que fueron contactadas.
- $w_{ra,k}$: : Factor de selección de la persona de primera y segunda fase

Una vez estimados los factores de ajuste, el ponderador de selección corregido por no contacto $w_{ra,k}^{NC}$ se expresa según la ecuación (15):

$$w_{ra,k}^{NC} = \hat{R}_{g,c} \cdot w_{ra,k} \quad (15)$$

Este ponderador se asigna a las 20.535 personas contactadas. A las personas no contactadas se asigna un valor blanco ("missing").

La Tabla 10 presenta la distribución de los veintiles y el promedio de la probabilidad de contacto de cada uno de ellos.

Tabla 10. Veintiles de ajuste por no contacto EBS 2025

Veintil	Personas seleccionadas	Personas contactadas	Personas no contactadas	Promedio probabilidad de contacto
Total	25.680	20.535	5.145	
1	1.284	568	716	0,103
2	1.284	606	678	0,163
3	1.284	625	659	0,207
4	1.284	681	603	0,249
5	1.284	737	547	0,293
6	1.284	811	473	0,344
7	1.284	935	349	0,403
8	1.284	1.022	262	0,474
9	1.284	1.076	208	0,549
10	1.284	1.146	138	0,621
11	1.284	1.145	139	0,685
12	1.284	1.168	116	0,740
13	1.284	1.185	99	0,788
14	1.284	1.229	55	0,832
15	1.284	1.245	39	0,871
16	1.284	1.265	19	0,903
17	1.284	1.267	17	0,929
18	1.284	1.269	15	0,950
19	1.284	1.275	9	0,966
20	1.284	1.280	4	0,983

Fuente: Instituto Nacional de Estadísticas.

5.3.2. Ajuste por no elegibilidad

A partir de esta etapa no se consideran aquellas personas que no forman parte de la población objetivo de la encuesta, es decir, personas consideradas como no elegibles por no residir en la dirección seleccionada al momento de la entrevista o nunca haber residido en dicha dirección, entre otras razones de no elegibilidad.

Este ajuste consiste en asignar un valor blanco (“missing”) en el ponderador ajustado por no contacto a aquellas personas clasificadas como no elegibles.

5.3.3. Ajuste por no respuesta

En esta etapa solo son consideradas las personas contactadas que resultaron elegibles, pues son estas las que cumplen con las características necesarias para participar en la encuesta. Sin embargo, es posible que algunas de estas personas no deseen participar o no cuenten con las condiciones para responder la encuesta, como es el caso, por ejemplo, de aquellos que se encuentran impedidos física o mentalmente al momento de la entrevista. En EBS 2025, de las 24.133 personas que resultaron elegibles (incluyendo las personas no contactadas), 13.600 fueron entrevistadas lo que significa que la tasa de respuesta es de 56,4%.

Si en la construcción del factor de expansión no se considera el fenómeno de la no respuesta a la unidad, las estimaciones serán representativas solamente de aquellas unidades que participan en la encuesta y, por lo tanto, una parte de la población no sería cubierta. Además, existe la posibilidad de que las estimaciones sean sesgadas en caso de que exista alguna relación entre la no respuesta y la variable de interés.

Para minimizar los problemas ocasionados por la falta de respuesta se utiliza el método de propensity score donde se modela la probabilidad de respuesta a partir de una regresión logística con variables obtenidas de Casen 2024 y de la hoja de ruta de EBS 2025. Una vez obtenida la probabilidad de respuesta, las personas se ordenan de menor a mayor probabilidad y son agrupadas en veintiles (20 grupos de igual tamaño).

En la tabla 11 se presentan las variables utilizadas en el modelo.

Tabla 11. Variables utilizadas en el modelo de no respuesta EBS 2025

Variable	Fuente
Proporción de encuestas logradas sobre el total de encuestas asignadas a la persona recolectora	Hoja de ruta EBS 2025
Contacto durante el fin de semana	Hoja de ruta EBS 2025
Número de llamados (logaritmo)	Hoja de ruta EBS 2025
Región	Casen 2024
Informante	Casen 2024
Hogar carente en seguridad	Casen 2024
Vivienda a menos de 8 cuadras o 1 km del transporte público	Casen 2024
Asiste a un establecimiento educacional	Casen 2024
Edad (al cuadrado)	Casen 2024
Ha sido discriminado en los últimos 12 meses	Casen 2024
50% de los integrantes del hogar se encuentran inactivos	Casen 2024
Acceso a internet	Casen 2024
Condición permanente o larga duración: Dificultad mental o intelectual	Casen 2024
Dificultad para comunicarse, comprender o ser comprendido por otros	Casen 2024



Variable	Fuente
Participación en alguna organización social, política o corporativa	Casen 2024
Dificultad para hacer o recibir llamadas	Casen 2024
Dependencia funcional	Casen 2024

Fuente: Instituto Nacional de Estadísticas.

Se espera que, al interior de las 20 celdas de ajuste, las personas sean similares en sus características y así, los entrevistados pueden representar a los no entrevistados. Bajo este supuesto, los ponderadores asociados a las personas contactadas elegibles que no respondieron deben ser redistribuidos entre las personas que sí respondieron. Para estos efectos, se debe calcular una razón de ajuste.

La determinación de la razón de ajuste viene dada por la ecuación (16) y se realiza al interior de cada veintil utilizando el ponderador ajustado por elegibilidad.

El numerador corresponde a la estimación del total de personas elegibles, y el denominador, a la estimación del total de personas que responden.

$$\hat{R}_{v,R} = \frac{\sum_{i \in \Phi_v} \sum_{j \in \theta_v} w_{ra,k,E}}{\sum_{i \in \Phi_v} \sum_{j \in \theta_{v,R}} w_{ra,k,R}} \quad (16)$$

Donde,

- v : : Índice del grupo para la corrección de no respuesta (veintil).
- Φ_v : : Conjunto de personas asignadas al veintil v .
- θ_v : : Conjunto de personas pertenecientes al veintil v elegibles y contactadas.
- $\theta_{v,R}$: : Conjunto de personas pertenecientes al veintil v que respondieron.
- $w_{ra,k,E}$: : Estimación del total de personas elegibles
- $w_{ra,k,R}$: : Estimación del total de personas que responde

Una vez estimados los factores de ajuste, el ponderador de selección corregido por no respuesta $w_{ra,k}^{NR}$ se expresa según la ecuación (17):

$$w_{ra,k}^{NR} = \hat{R}_{v,R} \cdot w_{ra,k,E} \quad (17)$$

Este ponderador se asigna a las 13.600 personas que respondieron y a las personas elegibles y no entrevistadas se les asigna un valor blanco ("missing").

La Tabla 12 presenta la distribución de los veintiles y el promedio de la probabilidad respuesta en cada uno de ellos.

Tabla 12. Veintiles de ajuste por no respuesta EBS 2025

Veintil	Personas elegibles y contactadas	Personas contactadas y no entrevistadas	Personas entrevistadas	Promedio probabilidad de respuesta
Total	18.988	5.388	13.600	
1	950	767	183	0,043
2	950	694	256	0,099
3	950	596	354	0,151
4	950	565	385	0,209
5	950	531	419	0,271
6	950	463	487	0,336
7	950	404	546	0,403
8	950	339	611	0,476
9	949	276	673	0,548
10	949	222	727	0,616
11	949	180	769	0,678
12	949	119	830	0,735
13	949	87	862	0,786
14	949	73	876	0,830
15	949	28	921	0,867
16	949	15	934	0,897
17	949	11	938	0,922
18	949	7	942	0,941
19	949	8	941	0,960
20	949	3	946	0,979

Fuente: Instituto Nacional de Estadísticas.

5.4. Suavizamiento

El suavizamiento es una técnica que permite reducir la cantidad de factores de expansión que se consideran que escapan de ciertos límites¹⁵, conocidos como valores atípicos y, para ello, se aplica una metodología mixta que incluye el los métodos de contracción a la media (CM) y R-K.

5.4.1. Método de contracción a la media (CM)

Este método tiene por objetivo que la distribución de los factores sea más homogénea y para ello contrae los factores de expansión, acercándolos al valor promedio de éstos. El factor suavizado o contraído a la media (o promedio) viene dado por:

$$w_i^s = \beta \cdot w_i + (1 - \beta) \cdot \bar{w}; 0 \leq \beta \leq 1 \quad (18)$$

¹⁵ Ver nota al pie n°15.



Donde,

- β : Es un parámetro real que permite controlar la intensidad de la contracción
 w_i : Factor ajustado por no respuesta
 $\bar{w}$: Media de los factores a nivel de estrato región-área y de valores atípicos y no atípicos

5.4.2. Método R-K

Este método da un tratamiento diferenciado a los conjuntos de datos atípicos o no atípicos. Por un lado, genera un aumento proporcional en los factores no atípicos y, por otro lado, produce una disminución proporcional en los factores atípicos. Esto significa que el método R-K disminuye en una proporción R los valores atípicos y aumenta en una proporción K los valores no atípicos como se indica en la figura 2.

Figura 2. Ordenamiento de factores de expansión atípicos y no atípicos

Factores		Aumento o disminución	
W_1	Atípicos	$W_1 \cdot (1 - r)$	$t_1 = \left(\sum_{i=1}^n W_i \right) - \left(\sum_{i=1}^n W_i \right) \cdot r$
W_2		$W_2 \cdot (1 - r)$	
:		:	
:		:	
W_n		$W_n \cdot (1 - r)$	
w_1	No atípicos	$w_1 \cdot (1 + k)$	$t_2 = \left(\sum_{j=1}^m w_j \right) + \left(\sum_{j=1}^m w_j \right) \cdot k$
w_2		$w_2 \cdot (1 + k)$	
:		:	
:		:	
w_m		$w_m \cdot (1 + k)$	

Fuente: Instituto

Nacional de Estadísticas.

5.4.3. Método mixto

El método mixto combina los métodos CM y R-K, y consiste en separar los factores de expansión en un subconjunto de factores no atípicos y otro de factores atípicos¹⁶. Para los factores no atípicos se aplica el método de contracción a la media considerando un β de 0,8 mientras que, para los factores atípicos el β es de 0,2 con el fin de que, en caso de tener un único valor atípico al interior del estrato, el método CM no tenga efecto sobre ese valor y se conserve el valor del factor atípico.

¹⁶ Para identificar los factores atípicos se utiliza como criterio el percentil de distribución del factor de expansión, para ello se define el valor $h = 1,5(p_{75} - p_{25})$ donde los valores con un valor inferior a $(p_{25} - h)$ o superior a $(p_{75} + h)$ se consideran atípicos.

Luego, se aplica el método R-K donde se unen las distribuciones tanto de factores atípicos como en un continuo que se puede lograr con valores de λ cercanos a la unidad. En este caso se utiliza un $\lambda=0,99$ para conservar una pequeña distancia entre el mayor de los factores no atípicos y el menor de los factores atípicos, después del suavizamiento.

5.5. Calibración de factores de expansión

La calibración de los factores de expansión sirve para ajustar los factores a los totales poblacionales dados por las proyecciones de población. Para EBS 2025 las proyecciones de población utilizadas son al 30 de agosto de 2025, fecha que coincide con la mitad del periodo de levantamiento de la encuesta.

Para la calibración se utiliza el método *Raking* que permite incorporar información auxiliar relacionadas con la variable de interés de EBS y tiene dos propósitos, el primero es que la encuesta, a partir de los pesos calibrados, sea capaz de reproducir totales poblacionales conocidos; y, el segundo es alcanzar mejoras en la precisión de los parámetros de interés por medio de la obtención de estimaciones con un menor error de muestreo (Alvarado, Pizarro & Guarda, 2020).

Las variables utilizadas como marginales de calibración son:

- Población total por región y área.
- Población total por región y sexo.
- Población total por tramo etario ([18-29], [30-44], [45-59], [60-79], [80+] y sexo.

La Tabla 13 muestra las proyecciones de población de 18 años y más, según región y área, y las proyecciones ajustadas a la población objetivo de la encuesta, es decir, población residente en viviendas particulares.

Tabla 13. Proyecciones de población al 30 de agosto 2025 según región y área

Región	Estimación de población total			Estimación de población residente en viviendas particulares		
	Urbano	Rural	Total	Urbano	Rural	Total
Nacional	13.979.194	1.801.364	15.780.558	13.975.021	1.800.623	15.775.644
Arica y Parinacota	185.021	16.727	201.748	184.705	16.464	201.169
Tarapacá	287.572	17.172	304.744	287.541	17.158	304.699
Antofagasta	529.770	21.124	550.894	529.303	21.053	550.356
Atacama	218.965	21.422	240.387	218.939	21.387	240.326
Coquimbo	559.802	125.101	684.903	559.655	125.062	684.717
Valparaíso	1.482.123	130.932	1.613.055	1.481.516	130.894	1.612.410
Metropolitana	6.411.737	239.120	6.650.857	6.410.187	239.030	6.649.217
O'Higgins	604.457	200.385	804.842	604.410	200.339	804.749
Maule	683.004	235.842	918.846	682.836	235.808	918.644



Región	Estimación de población total			Estimación de población residente en viviendas particulares		
	Urbano	Rural	Total	Urbano	Rural	Total
Ñuble	291.973	121.086	413.059	291.943	121.085	413.028
Biobío	1.178.476	145.713	1.324.189	1.178.225	145.706	1.323.931
La Araucanía	576.172	224.652	800.824	575.972	224.648	800.620
Los Ríos	237.156	88.160	325.316	237.056	88.145	325.201
Los Lagos	531.775	185.633	717.408	531.608	185.612	717.220
Aysén	66.611	16.147	82.758	66.581	16.102	82.683
Magallanes	134.580	12.148	146.728	134.544	12.130	146.674

Fuente: Instituto Nacional de Estadísticas.

Para obtener la proyección de población residente en viviendas particulares se utiliza la proporción de población residente en viviendas particulares sobre el total de población según los resultados del Censo de Población y Vivienda 2017 a nivel de cada una de las marginales de calibración.

El resultado del proceso de calibración se presenta en la tabla 14 donde se observa que, luego del ajuste, los factores de expansión alcanzan un total de 15.775.644 personas de 18 años y más.

Tabla 14. Distribución de la muestra y factor de expansión calibrado por región y área, EBS 2025

Región	Muestra lograda	Factor de expansión calibrado a población de 18 años y más
Nacional	13.600	15.775.644
Urbano	11.327	13.975.021
Rural	2.273	1.800.623
Arica y Parinacota	645	201.169
Tarapacá	695	304.699
Antofagasta	741	550.356
Atacama	671	240.326
Coquimbo	841	684.717
Valparaíso	1.034	1.612.410
Metropolitana	1.738	6.649.217
O'Higgins	884	804.749
Maule	909	918.644
Ñuble	748	413.028
Biobío	1.006	1.323.931
La Araucanía	902	800.620
Los Ríos	772	325.201
Los Lagos	873	717.220
Aysén	549	82.683
Magallanes	592	146.674

Fuente: Instituto Nacional de Estadísticas.

5.6. Redondeo Probabilístico

Finalmente, como los resultados son referidos a personas, los valores de los factores de expansión deben expresarse en números enteros. Para ello, se aplica un redondeo probabilístico cuya ventaja es que los nuevos pesos calibrados no produzcan un sesgo en las estimaciones de interés.

El redondeo probabilístico distribuye aleatoriamente la parte decimal de los factores, de acuerdo con una distribución de probabilidades según el valor de esta y utilizando semillas de aleatorización que permiten replicar el procedimiento y alcanzar el stock poblacional a un nivel deseado que, en caso de EBS, corresponde a nivel nacional.

Para llevar a cabo este procedimiento se utiliza el método *sample cube*, que trabaja con submuestras balanceadas considerando los stocks poblacionales, dados por las marginales de calibración: región – área, región – sexo y grupo edad – sexo. Esto permite asignar valores 0 y 1, obteniéndose factores de expansión enteros y consistentes con el total nacional según la proyección de población.

Es importante recalcar que este método no garantiza el ajuste exacto para cada una de las categorías de las marginales ya que solo es posible aproximarse, en este caso, al total nacional. Esto se consigue a través del uso de semillas de aleatorización donde se evaluaron 151 semillas (entre 500 y 650) seleccionando aquella que presenta la menor diferencia respecto a las proyecciones de población utilizadas.

Una vez finalizado el proceso de redondeo de factores de expansión, se verifica que estos cumplan con ciertos criterios. En primer lugar, se verifica que la suma total de los factores de expansión redondeados alcance exactamente el stock poblacional de 15.775.644 personas de 18 años y más residentes en viviendas particulares, según información del INE.

Adicionalmente, se valida que la diferencia entre la suma de los factores redondeados al nivel de cada marginal y los totales definidos para cada una de ellas sea la menor posible. El cumplimiento de este criterio permite afirmar que los factores de expansión redondeados reflejan la estructura poblacional y no se introducen sesgos en el estimador.

6. CÁLCULO Y EVALUACIÓN DE LAS ESTIMACIONES

Este capítulo tiene como objetivo presentar los métodos de estimación de varianza en diseños muestrales complejos que, en el caso de EBS, corresponde a un diseño bifásico elaborado a partir de Casen 2024. Para esto, se consideran las variables insatisfacción con la vida y balance afectivo negativo, sobre las cuales se obtienen estimaciones de proporciones junto con sus errores muestrales asociados para los dominios de estudio definidos.

6.1. Estimación de la varianza compleja

Los métodos exactos de estimación de la varianza en diseños de muestras estándares, cuando son aplicables, son la mejor forma de estimarla. Sin embargo, los diseños muestrales empleados en encuestas sociales, como es el caso de EBS, son más complejos que el muestreo aleatorio simple o estratificado.

Un diseño de muestreo complejo se origina cuando existe más de una etapa de selección. En el caso de EBS, el diseño muestral se ancla a la encuesta Casen 2024¹⁷, donde la selección de las unidades muestrales se realiza a partir de la muestra lograda en dicha encuesta que ya posee un diseño complejo, lo que añade otro grado de complejidad al cálculo de precisión de las estimaciones debido a la varianza adicional que implica el submuestreo de segunda fase (Gutiérrez, 2015).

En el muestreo estratificado la estimación de la varianza total se obtiene mediante la suma de las varianzas estimadas en cada estrato, y la varianza en cada estrato se estima a partir de la variabilidad de la característica de interés entre los conglomerados. En este punto cabe señalar que EBS cuenta con un diseño estratificado donde los conglomerados no corresponden a una etapa en la selección, pero se encuentran presentes en su diseño ya que estos se obtienen desde Casen 2024.

Una aproximación para la estimación de varianza comúnmente utilizada en encuestas complejas, de muestreo polietápico por conglomerados, corresponde al método de estimación de varianza por conglomerado último (EVCU) o por reemplazo (WR por sus siglas en inglés). Este método permite capturar cierto grado de covarianza por medio de la correlación intraclásica de las unidades al interior de los conglomerados, aun cuando las estimaciones se basen en muestras sin traslape. Asimismo, es posible la captura de covarianza a través de la correlación intraclásica de las unidades al interior de los estratos o dominios comunes (INE, 2022).

Estas estimaciones con EVCU, pueden efectuarse indistintamente por linealización mediante el desarrollo en Series de Taylor o con una técnica de replicación. Con este enfoque, las UPM se tratan como si se hubieran

¹⁷ Casen 2024 presenta un diseño muestral probabilístico, estratificado y bietápico.

seleccionado con reemplazo al interior de los estratos. En tal caso, cada elemento o individuo solo tiene que identificarse por el estrato y por la UPM (dentro del estrato) a la que pertenezca.

Para efectos de la estimación de varianzas no se necesita información sobre las etapas de muestreo por debajo del nivel de las UPM. De este modo la descripción del plan de muestreo para la estimación de varianzas, se simplifica de forma que se asemeja a un muestreo por conglomerados estratificados en una etapa, esto es, una muestra estratificada de conglomerados finales completamente enumerados. Este enfoque por conglomerados finales produce una buena aproximación para la estimación de la varianza, siempre que sea razonable partir de la hipótesis de una primera etapa con reemplazo.

Así, cuando el plan de muestreo se describe como EVCU, solo se necesitan tres variables del diseño de la encuesta para la estimación de varianzas:

1. La variable de ponderación de la muestra o factor de expansión designado “fexp” donde el nivel de expansión territorial es regional.
2. La variable de estratificación “estrato_ebs” correspondiente a la combinación de región-área en EBS 2025.
3. La variable UPM (o pseudo-conglomerado) designada como “varunit” y obtenida desde Casen 2024.

El empleo de EVCU para la estimación de varianzas considerando el diseño complejo puede dar lugar a una ligera sobreestimación. Sin embargo, al realizar el análisis en general, esta sobreestimación se asume aceptable considerando la relativa simplicidad de la aproximación mediante EVCU. No obstante, hay que tener en cuenta que la sobreestimación puede ser apreciable si hay varios estratos en los que el muestreo en la primera etapa se hace sin reemplazo y con grandes fracciones de muestreo. En este caso, puede ser preferible optar por un software que cuente con la opción de incorporar los factores de corrección por población finita *cpf* o corrección por finitud para la primera etapa.

6.2. Variables que identifican el diseño muestral complejo

En la muestra de EBS 2025, para identificar los estratos definidos en el diseño se utiliza la variable “estrato_ebs”. Por otra parte, la selección de personas se realizó directamente al interior de cada estrato, sin la necesidad de considerar primero a los conglomerados. Sin embargo, el diseño de la primera fase (Casen 2024), sí contempla la utilización de conglomerados para la selección de unidades, por lo que, al momento de realizar las estimaciones, es necesario considerarlos en la definición del diseño complejo a través de la variable “varunit” obtenida desde Casen 2024. Según lo anterior, en total, en la muestra lograda, las personas se distribuyen en 32 estratos y 8.559 conglomerados.

6.3. Estadísticos asociados a los parámetros de interés

A partir de la muestra lograda en EBS 2025, se obtienen los parámetros asociados a las variables de interés y con ello, sus principales estadísticos. Estos parámetros corresponden a las proporciones de las siguientes variables:

- Insatisfacción con la vida (a1) correspondiente al porcentaje de personas con la categoría “Totalmente insatisfecho” o “Insatisfecho” con su vida.
- Balance afectivo negativo (affective_sw) que corresponde al porcentaje de personas que presentan un balance negativo, es decir, las emociones negativas (enojado(a), triste) superan las positivas (tranquilo(a), feliz)¹⁸.

Las tablas 15 y 16 presentan la estimación y errores de la variable insatisfacción con la vida y balance afectivo negativo, respectivamente, a nivel nacional, por área y a nivel regional.

Tabla 15. Errores muestrales por región y área para la variable Insatisfacción con la vida, EBS 2025

Región	Insatisfacción con la vida		
	Estimación	Error absoluto	Error relativo
Nacional	8,1%	0,7%	8,0%
Urbano	8,2%	0,7%	8,7%
Rural	8,0%	1,5%	18,5%
Arica y Parinacota	7,2%	2,1%	29,1%
Tarapacá	7,8%	2,2%	28,4%
Antofagasta	6,7%	2,1%	31,9%
Atacama	6,9%	1,9%	28,3%
Coquimbo	6,7%	1,8%	26,9%
Valparaíso	11,5%	2,2%	19,3%
Metropolitana	7,5%	1,3%	16,9%
O'Higgins	8,4%	1,9%	22,5%
Maule	7,4%	1,8%	24,6%
Ñuble	9,3%	2,4%	25,9%
Biobío	7,8%	1,9%	24,2%
La Araucanía	10,0%	2,1%	21,3%
Los Ríos	8,1%	2,0%	25,0%
Los Lagos	8,7%	2,1%	23,7%
Aysén	7,0%	2,1%	30,2%
Magallanes	5,7%	2,2%	37,8%

¹⁸ El balance negativo se construye considerando las personas que presentan valores negativos en la variable affective_sw.

Fuente: Instituto Nacional de Estadísticas.

Tabla 16. Errores muestrales por región y área para la variable balance afectivo negativo, EBS 2023

Región	Balance afectivo negativo		
	Estimación	Error absoluto	Error relativo
Nacional	8,4%	0,7%	8,5%
Urbano	8,6%	0,8%	9,1%
Rural	6,9%	1,4%	20,0%
Arica y Parinacota	7,4%	2,1%	27,8%
Tarapacá	9,0%	2,4%	26,5%
Antofagasta	5,4%	1,8%	34,1%
Atacama	6,9%	2,2%	31,3%
Coquimbo	8,7%	2,1%	24,4%
Valparaíso	10,6%	2,1%	20,0%
Metropolitana	8,3%	1,4%	17,5%
O'Higgins	10,3%	2,1%	20,1%
Maule	6,8%	1,8%	26,3%
Ñuble	6,9%	1,9%	28,2%
Biobío	8,9%	2,0%	21,9%
La Araucanía	8,7%	2,0%	22,9%
Los Ríos	7,4%	2,0%	26,7%
Los Lagos	9,0%	2,2%	24,3%
Aysén	3,8%	1,8%	47,3%
Magallanes	5,0%	1,8%	36,2%

Fuente: Instituto Nacional de Estadísticas.



7. ANEXOS

Anexo N°1: Categoría elegibilidad y códigos de disposición final (CDF) EBS 2025

Categoría elegibilidad	CDF	Glosa CDF
Elegible entrevistada	11	Entrevista completa
	12	Entrevista parcial
Elegible rechazada	22	Rechazo persona informante
Elegible no entrevistada por otros motivos	41	Persona informante impedida física y/o mentalmente para contestar
	42	Persona informante no habla idioma español
	43	Muerte de la persona informante o algún(a) familiar
	46	Pérdida o errores en la transmisión de datos
	49	Otra razón de no entrevista
	81	Calidad de audio mala o intermitente
	82	Lugar o actividad de la persona informante no permite una entrevista
Elegible no contactada	33	Persona informante no ubicable, se contacta otra persona
	71	Marca tono, pero sin respuesta
	72	Responde buzón de voz
	73	Suena ocupado
	74	Se encuentra apagado o fuera del área de cobertura
	75	Número telefónico no existe
	76	Número telefónico equivocado
	79	Otra razón de no contacto
No elegible	61	Establecimiento, oficina de gobierno u otra organización
	62	Institución (por ejemplo: hospital, cárcel, residencias para personas mayores, etc.)
	64	Vivienda en demolición, incendiada, destruida o erradicada
	65	Vivienda desocupada (por ejemplo: en venta, arriendo, etc.)
	66	Vivienda de veraneo o de uso temporal
	69	Otra razón para ser considerada no elegible
	91	Persona informante/hogar ya no reside en la dirección seleccionada
	92	Persona informante/hogar nunca residió en la dirección seleccionada

Anexo N°2: Sintaxis en R para la estimación de las variables de interés

1. Cargar base y crear diseño para estimación

```
library(openxlsx)
library(haven)
library(tidyverse)
library(survey)
library(sjlabelled)

rm(list = ls())
options(scipen = 999)

ebs2025 <- read_dta("Insumos/ebs2025_analisis.dta")

ebs2025 <- ebs2025 %>%
  mutate(insatis = as.factor(ifelse(a1 %in% c(1:2),1,0)),
         balance_neg = as.factor(case_when(affective_sw < 0 ~ 1,
                                           affective_sw %in% c(0:8) ~ 0,
                                           .default = NA)),
         Nacional = 1)
```

1.1. Diseño para estimación con diseño original EBS (1 fase) y conglomerado Casen:

```
d_1f_upm <- svydesign(ids= ~varunit, #varunit desde Casen 2024
                   strata = ~estrato_ebs, #region-área
                   weights = ~fexp,
                   data= ebs2025)
```

2. Estimación variable insatisfacción con la vida

2.1. Estimación proporción

```
est_tasa <- function(dominio,variable){
  formula_var <- as.formula(paste0("~",variable))
  formula_dom <- as.formula(paste0("~",dominio))

  estimacion <- svyby(formula_var,formula_dom, d_1f_upm, na.rm = T, level = 0.95,
                     vartype = c("se","cv"), svymean)
  return(estimacion)
}

est_nac <- est_tasa("Nacional","insatis") %>%
  rename(region = Nacional) %>%
  mutate(region = "Nacional")

est_reg <- est_tasa("region","insatis")

est_area <- est_tasa("area","insatis") %>%
  rename(region = area) %>%
  mutate(region = ifelse(region == 1, "Urbano","Rural"))
```

2.2. Grados de libertad por dominio

```
#Nacional
gl_nac <- data.frame(region = "Nacional", gl = degf(d_1f_upm))

#Regional
gl_reg <- 1:16 %>% lapply(function(r) {
  design <- subset(d_1f_upm, region == r)
  data.frame(region = as.factor(r), gl = degf(design))) %>%
  bind_rows()

#Área
gl_area <- data.frame(region = c("Urbano", "Rural"),
  gl = sapply(1:2, function(a) degf(subset(d_1f_upm, area == a))))
```

2.3. Unir resultados

```
prop_insatis <- rbind(est_nac, est_area, est_reg) %>%
  select(1,3,5,7)

gl_insatis <- rbind(gl_nac, gl_area, gl_reg)

tabla_insatis <- left_join(prop_insatis, gl_insatis, by = "region")
```

2.4. Totales muestra

```
n_reg_insatis <- ebs2025 %>%
  group_by(region) %>%
  summarise(n = n(),
    n_insatis = sum(insatis == 1))

n_area_insatis <- ebs2025 %>%
  group_by(area) %>%
  summarise(n = n(),
    n_insatis = sum(insatis == 1)) %>%
  rename(region = area) %>%
  mutate(region = ifelse(region == 1, "Urbano", "Rural"))

region <- "Nacional"
n <- sum(n_reg_insatis$n)
n_insatis <- sum(n_reg_insatis$n_insatis)
n_nac_insatis <- data.frame(region, n, n_insatis)

n_insatis <- rbind(n_nac_insatis, n_area_insatis, n_reg_insatis)

tabla_insatis <- left_join(tabla_insatis, n_insatis, by = "region")
```

2.5. Calcular errores y aplicar estándar de calidad

```
tabla_insatis <- tabla_insatis %>%
  rename(SE = se.insatis1) %>%
  mutate(Ts = qt(0.975, gl),
    error_abs = Ts * SE,
    error_rel = error_abs/insatis1,
```

```

se_max = ifelse(insatis1 < 0.5,
                (insatis1^(2/3))/9,
                ((1-insatis1)^(2/3)/9)),
est_calidad = case_when(n >= 60 & gl >= 9 & SE < se_max ~ "fiable",
                        n >= 60 & gl >= 9 & SE > se_max ~ "poco fiable",
                        n < 60 | gl < 9 ~ "no fiable",
                        TRUE ~ NA),
region = factor(region, levels = c("Nacional", "Urbano", "Rural", 15, 1:5, 13, 6, 7, 16, 8,
9, 14, 10:12))) %>%
  arrange(region)

rm(est_area, est_nac, est_reg, gl_insatis, gl_area, gl_nac, gl_reg, n_insatis, n_area_insatis, n_nac_
insatis, prop_insatis, n_reg_insatis)

```

3. Estimación variable balance afectivo negativo

3.1. Estimación proporción

```

est_nac <- est_tasa("Nacional", "balance_neg") %>%
  rename(region = Nacional) %>%
  mutate(region = "Nacional")

est_reg <- est_tasa("region", "balance_neg")

est_area <- est_tasa("area", "balance_neg") %>%
  rename(region = area) %>%
  mutate(region = ifelse(region == 1, "Urbano", "Rural"))

```

3.2. Grados de libertad por dominio

```

#Nacional
gl_nac <- data.frame(region = "Nacional", gl = degf(d_1f_upm))

#Regional
gl_reg <- 1:16 %>% lapply(function(r) {
  design <- subset(d_1f_upm, region == r)
  data.frame(region = as.factor(r), gl = degf(design))) %>%
  bind_rows()

#Área
gl_area <- data.frame(region = c("Urbano", "Rural"),
  gl = sapply(1:2, function(a) degf(subset(d_1f_upm, area == a))))

```

3.3. Unir resultados

```

prop_balance_neg <- rbind(est_nac, est_area, est_reg) %>%
  select(1, 3, 5, 7)

gl_balance_neg <- rbind(gl_nac, gl_area, gl_reg)

tabla_balance_neg <- left_join(prop_balance_neg, gl_balance_neg, by = "region")

```

3.4. Totales muestra

```

n_reg_balance_neg <- ebs2025 %>%
  group_by(region) %>%
  summarise(n = n(),
            n_balance_neg = sum(balance_neg == 1, na.rm = T))

n_area_balance_neg <- ebs2025 %>%
  filter(balance_neg == 1) %>%
  group_by(area) %>%
  summarise(n = n(),
            n_balance_neg = sum(balance_neg == 1, na.rm = T)) %>%
  rename(region = area) %>%
  mutate(region = ifelse(region == 1, "Urbano", "Rural"))

region <- "Nacional"
n <- sum(n_reg_balance_neg$n)
n_balance_neg <- sum(n_reg_balance_neg$n_balance_neg)
n_nac_balance_neg <- data.frame(region, n, n_balance_neg)

n_balance_neg <- rbind(n_nac_balance_neg, n_area_balance_neg, n_reg_balance_neg)

tabla_balance_neg <- left_join(tabla_balance_neg, n_balance_neg, by = "region")

```

3.5. Calcular errores y aplicar estándar de calidad

```

tabla_balance_neg <- tabla_balance_neg %>%
  rename(SE = se.balance_neg1) %>%
  mutate(Ts = qt(0.975, gl),
         error_abs = Ts * SE,
         error_rel = error_abs / balance_neg1,
         se_max = ifelse(balance_neg1 < 0.5,
                          (balance_neg1^(2/3))/9,
                          ((1 - balance_neg1)^(2/3))/9),
         est_calidad = case_when(n >= 60 & gl >= 9 & SE < se_max ~ "fiable",
                                  n >= 60 & gl >= 9 & SE > se_max ~ "poco fiable",
                                  n < 60 | gl < 9 ~ "no fiable",
                                  TRUE ~ NA),
         region = factor(region, levels = c("Nacional", "Urbano", "Rural", 15, 1:5, 13, 6, 7, 16, 8,
                                             9, 14, 10:12))) %>%
  arrange(region)

```

4. Guardar resultados

```

resultados <- list()

resultados[["Insatisfacción"]] <- tabla_insatis
resultados[["Balance afectivo"]] <- tabla_balance_neg

wb <- createWorkbook()

for (i in seq_along(resultados)) {
  sheet_name <- names(resultados)[i]
  addWorksheet(wb, sheet_name)

  # Escribir datos

```

```
writeData(wb, sheet_name, resultados[[i]],
          headerStyle = createStyle(textDecoration = "bold", fgFill = "#98dfaf", halign =
"center"))

# Formato de números y porcentajes
pct_style <- createStyle(numFmt = "0.0%")
num_style <- createStyle(numFmt = "#,##0")
dec_style <- createStyle(numFmt = "#,##0.00")

# Aplicar estilo diferenciando columnas
addStyle(wb, sheet = sheet_name, style = pct_style,
         cols = c(2:4,9:11), rows = 2:(nrow(resultados[[i]]) + 1), gridExpand = T, stack
= T)

addStyle(wb, sheet = sheet_name, style = num_style,
         cols = c(5:7), rows = 2:(nrow(resultados[[i]]) + 1), gridExpand = T, stack = T)

addStyle(wb, sheet = sheet_name, style = dec_style,
         cols = c(8), rows = 2:(nrow(resultados[[i]]) + 1), gridExpand = T, stack = T)
}

xlsx <- paste0("Resultados/",paste0(gsub("-", "", Sys.Date()), "_Estimaciones EBS 2025.xlsx"))
saveWorkbook(wb, xlsx, overwrite = TRUE)
```



8. REFERENCIAS

Cochran, W. (1998). *Técnicas de Muestreo*. México D.F.: Compañía Editorial Continental, S.A. de C.V.

Heeringa, S. G., West, B. T., & Berglund, P. A. (2010). *Applied Survey Data Analysis*. USA: Chapman & Hall / CRC Press.

INE. (marzo de 2020). Fundamentos del Estándar para la evaluación de la calidad de las estimaciones en encuestas de hogares. Obtenido de: <https://www.ine.gob.cl/inicio/documentos-de-trabajo/documento/fundamentos-del-est%C3%A1ndar-para-la-evaluaci%C3%B3n-de-la-calidad-de-las-estimaciones-en-encuestas-de-hogares>

INE, MDSF (2024). Metodología de Diseño Muestral EBS 2023. Obtenido de: https://observatorio.ministeriodesarrollosocial.gob.cl/storage/docs/bienestar-social/2023/Metodologia_Disenio_Muestral_EBS_2023.pdf

Kish, L. (1965). *Survey Sampling*. Nueva York: John Wiley & Sons.



