



**Encuesta
Nacional de
Salud
2016 -2017**

ANÁLISIS ESTADÍSTICO



SUBSECRETARÍA DE SALUD PÚBLICA
DIVISIÓN DE PLANIFICACIÓN SANITARIA
Departamento de Epidemiología

5. ANÁLISIS ESTADÍSTICO ENS 2016-2017

5.1. Cálculo para estimadores puntuales y por intervalos de confianza de proporciones y medias

La estimación de frecuencias simples, prevalencias y promedios poblacionales se llevó a cabo respetando la metodología del muestreo complejo y utilizando los factores de expansión correspondientes (se utilizó el módulo de muestras complejas para los análisis). Se calcularon medias y proporciones con sus respectivos errores estándar e intervalos de confianza al 95% utilizando el método de linealización de Taylor (Wolter, 1985), sobre la base del diseño de muestreo complejo. Se utilizó este método dado que permite la comparabilidad entre las versiones de la ENS. Este método obtiene una aproximación lineal para los estimadores y utiliza la varianza estimada para esta aproximación para calcular la varianza de los estimadores. Todos los software disponibles en nuestro medio y que incluyen módulos especializados para análisis de muestras complejas (SAS; STATA; R; SPSS) permiten setear la elección de la linealización de Taylor como método de análisis (SPSS la tiene por default).

Para comparar prevalencias ajustando por edad, sexo y nivel educacional, se construyeron modelos lineales multivariados generalizados (logísticos para eventos binomiales y lineales para eventos numéricos), utilizando el módulo de muestras complejas y factores de expansión correspondientes, con todas las variables de interés¹. De los coeficientes del modelo se derivaron las razones de chance (OR). Los errores estándar de los OR se estimaron también utilizando linealización de Taylor.

Todos los procedimientos fueron realizados utilizando el módulo para muestras complejas del software SPSS o STATA. Las sintaxis de F3 CIDI fueron aplicadas en SAS.

5.2. Consideraciones generales para el análisis e interpretación de resultados

A continuación se resumen algunas consideraciones metodológicas generales comunes y necesarias para la interpretación adecuada de los análisis realizados en ENS 2016-2017.

- a) Para el caso de variables categóricas, se presentan tablas de resultados con las prevalencias expandidas (los valores representan a la población de 15 y más años de Chile, a junio de 2015) , sus respectivos intervalos de confianza del 95% (IC95%) y los tamaños de población a los que expanden, desgregados por sexo, grupos de edad (15-24, 25-44, 45-64 y 65 y más), años totales de estudio cursados aprobados (<8 , 8-12, >12 años) y zona (urbano/rural). Se presentan además las prevalencias por región, excepto en el caso de resultados de submuestras que poseen tamaños muestrales pequeños que pueden afectar la precisión de una estimación regional. En el caso de las variables cuantitativas continuas, se presentan en tabla su media, mediana, percentiles 25 y 75, valores mínimos y máximos y su tamaño muestral.

En los informes publicados de ENS2016 las prevalencias fueron calculadas sin excluir valores atípicos, las excepciones están especificadas en los informes por tema.

Se ha priorizado el dejar los datos crudos para que cada investigador pueda rehacer los análisis en el futuro con sus propios criterios de validación y excluyendo o no valores atípicos o outlayers.

Dada la existencia de submuestras aleatorias al interior de la ENS, la expansión a nivel nacional ha sido debidamente cautelada, utilizando distintos factores de expansión para un mismo participante (poner link de factor de expansión)

- b) Los n muestrales varían de tabla en tabla según el denominador utilizado para las estimaciones. Se habla de prevalencias en “población general” cuando el denominador del análisis son todos los participantes de la ENS que disponen de esa medición. Otra fuente de variación de estos n es la no respuesta (valores perdidos) los que son cambiantes dependiendo del ítem. Se habla de distribución de frecuencias cuando el denominador es una “subpoblación de enfermos”, por ejemplo frecuencia de diabéticos con microalbuminuria.

Se usaron modelos multivariados para evaluar el efecto del sexo, la edad y los años totales de estudio aprobados sobre las variables en estudio. Para el caso de variable dependiente cuantitativa con distribución normal se usó regresión lineal y cuando la variable no cumplió con este supuesto, ésta se categorizó como binaria y se analizó a través de una regresión logística.

Las variables categóricas se transformaron también en dicotómicas y se analizaron a través de regresión logística. En las tablas se presentan los valores p asociados a cada variable y en algunos casos se presentan además los Odds Ratio (OR) e intervalos de confianza (I. C.) de 95%.

El análisis multivariado permitió análisis ajustados por sexo, edad y años de estudio aprobados para cada variable dependiente en estudio.

- c) En el análisis regional, se muestra un gráfico que describe las prevalencias según sexo y región. Se exhibe adicionalmente una tabla con las prevalencias regionales comparadas con la prevalencia de la XIII Región Metropolitana. En este último análisis se muestran Odds Ratios con I. C. de 95% ajustados por edad y sexo (regresión logística utilizando el módulo de muestras complejas).
- d) A pesar del diseño muestral representativo del nivel nacional y regional, no todas las estimaciones alcanzan poder estadístico suficiente para el análisis regional, especialmente aquellas prevalencias pequeñas y que provienen de submuestras aleatorias como es el caso de: la submuestra de exámenes EX2 n=3700 (lípidos, transaminasas, electrolitos urinarios, metales pesados...), la submuestra EX3 n=1300 (ioduria, TSH, TPO, T4L) o la submuestra F3 n=3500 (módulo ampliado de salud mental). Se debe tener en cuenta que estas tres submuestras aleatorias no fueron calibradas (postestratificadas en la expansión de datos) en 120 celdas de edad/zona/región/sexo como lo fueron la muestra F1, F2 y EX1. Sin embargo en el caso de la submuestra F3, se hizo un reajuste posterior del factor de expansión y se logró calibrar a las 14 celdas correspondiente a los totales del cruce de sexo y edad (18-24, 25-34, 35-44, 45-54, 55-64, 65-74 y 75 y más años) y a las 15 totales regionales, de manera independiente. Sin embargo, se debe tener cuidado en la interpretación de resultados regionales en este módulo cuando el coeficiente de variación es superior a 30% ya que se trata de estimaciones poco precisas y por lo tanto no confiables.

Se presentan en ENS entonces, análisis de indicadores que pueden tener distinto nivel de precisión, dado que sus prevalencias estimadas y sus tamaños muestrales son variables. Sin embargo, esto no descarta que alguna prevalencia de mayor tamaño tenga representatividad regional en una submuestra. Para un adecuado análisis de estos aspectos se recomienda mirar el error relativo y absoluto de las estimaciones en las distintas submuestras logradas al interior de ENS 2016-2017, se sugiere consultar la sección de diseño muestral donde se exhiben los errores muestrales para distintas magnitudes de prevalencia y diversos tamaños muestrales.

5.3. Plan de análisis y construcción de indicadores de salud

El plan de análisis fue consensuado con MINSAL y se generaron indicadores cuyas definiciones, operacionalizaciones, puntos de corte, factores de expansión y denominadores se describen en los informes y fichas técnicas.

ⁱ McCullagh, P., Nelder, J. A., & McCullagh, P. (1989). Generalized linear models (Vol. 2). London: Chapman and Hall.

Referencias Bibliográficas de los capítulos 3 y 4 (Diseño muestral y Factores de expansión)

Gaziano, Cecile (2005). "Comparative Analysis of Within-Household Respondent Selection Techniques." *Public Opinion Quarterly* 69:124-157.

Instituto Nacional de Estadísticas (2010). "Nueva encuesta nacional de empleo. Manual conceptual y metodológico. Diseño Muestral". Santiago, Chile.

Instituto Nacional de Estadísticas (2014a). "Actualización y Proyecciones Cortas de Población. País y Regiones Total: Actualización Población 2002-2012 y Proyecciones 2013-2020". Santiago, Chile.

Instituto Nacional de Estadísticas (2014b). "Actualización y Proyecciones Cortas de Población. País y Regiones por Áreas Urbana-Rural: Actualización Población 2002-2012 y Proyecciones 2013-2020". Santiago, Chile.

Kish, Leslie (1949). "A Procedure for Objective Respondent Selection within the Household", *Journal of the American Statistical Association* 44 (247): 380–387

Kolenikov, S. (2014). "Calibrating survey data using iterative proportional fitting (raking)". *The Stata Journal*, Number 1, pp. 22–59.

McCullagh, P., Nelder, J. A., & McCullagh, P. (1989). *Generalized linear models* (Vol. 2). London: Chapman and Hall.

Ministerio de Desarrollo Social (2012a). "Diseño Muestral y Cálculo de Factores de Expansión Encuesta de Caracterización Socioeconómica Nacional (Casen) 2011". Santiago, Chile.

Ministerio de Desarrollo Social (2012b). "Diseño Muestral y Cálculo de Factores de Expansión Encuesta de Actividades de Niños, Niñas y Adolescentes (EANNA) 2012". Santiago, Chile.

Ministerio de Salud (2007). "II Encuesta CVS 2006 II Encuesta de Calidad de Vida y Salud Chile 2006, Informe de Resultados, Total Nacional". Santiago, Chile.

Rosenbaum, P. R., Rubin, D. B. (1983). „The Central Role of the Propensity Score in Observational Studies for Causal Effects". *Biometrika*, Vol. 70, No. 1. (Apr., 1983), pp. 41-55.

Valliant, R. Dever, J.A., Kreuter, F. (2013). "Practical Tools for Designing and Weighting Survey Samples". New York: Springer.