

TESIS DOCTORAL

NUEVAS TÉCNICAS DE INVESTIGACIÓN POR MUESTREO
APLICADAS A ENCUESTAS DE SALUD:
CALIBRACIÓN DE ESTIMADORES

Andrés Cabrera León



DIRECTORES

PROF.^a DRA. D.^a María del Mar Rueda García / Universidad de Granada

PROF. DR. D. Antonio Daponte Codina / Escuela Andaluza de Salud Pública

PROF. DR. D. Alejandro Jadad / University of Toronto, University Health Network



Universidad de Granada

Departamento de Estadística e Investigación Operativa

Facultad de Ciencias

Universidad de Granada

Febrero 2016

Editor: Universidad de Granada. Tesis Doctorales
Autor: Andrés Cabrera León
ISBN: 978-84-9125-869-8
URI: <http://hdl.handle.net/10481/43669>

TESIS DOCTORAL

NUEVAS TÉCNICAS DE INVESTIGACIÓN POR MUESTREO APLICADAS A ENCUESTAS DE SALUD: CALIBRACIÓN DE ESTIMADORES

ANDRÉS CABRERA LEÓN

DIRECTORES

PROF.^a DRA. D.^a MARÍA DEL MAR RUEDA GARCÍA
Universidad de Granada

PROF. DR. D. ANTONIO DAPONTE CODINA
Escuela Andaluza de Salud Pública

PROF. DR. D. ALEJANDRO JADAD
University of Toronto, University Health Network



Universidad de Granada

Departamento de Estadística e Investigación Operativa
Facultad de Ciencias
Universidad de Granada
Febrero 2016

Andrés Cabrera León: *Nuevas técnicas de investigación por muestreo aplicadas a encuestas de salud: Calibración de estimadores*, Universidad de Granada; Febrero 2016

Profesor. Escuela Andaluza de Salud Pública

Investigador. Centro de Investigación Biomédica en Red en Epidemiología y Salud Pública (CIBERESP)

DIRECTORES:

Prof.^a Dra. D.^a María del Mar Rueda García

Prof. Dr. D. Antonio Daponte Codina

Prof. Dr. D. Alejandro Jadad



Esta obra está bajo una [Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/).

Se puede acceder en su versión electrónica más actualizada en [TESEO](https://teseo.es/).

Ohana significa «familia».
Familia significa «nadie se queda atrás, u olvidado».

— Lilo & Stitch

Dedicado a mi incondicional y amado Equipo A.

ABSTRACT

INTRODUCTION This doctoral thesis is entitled “*New sampling research techniques applied to health surveys: calibration of estimators*”.

One of the main requirements in scientific research is to achieve the maximum measurement accuracy. Variations in measurements are characteristic of medicine and other biological sciences, but there are additional sources of variation that result from the act of measuring itself, i.e., the *precision and validity*. Precision (also known as reliability, reproducibility, concordance, or agreement) is associated with *random error* and is therefore unpredictable and inevitable, but it is estimable. For its part, validity is associated with systematic error or *bias* and is predictable and correctable. Precision estimation and control are often related to *Statistics*, while validity estimation and control are often related to *Epidemiology*.

Relevant sampling research methodologies developed over the past 25 years that both minimize bias and increase the precision of estimations for variables measured through samples, include:

Randomized response models: these enhance anonymity protection and achieve superior reliability and higher response rates.

Multiple frames: these overcome problems caused by the lack of coverage in certain sampling frames, including different data gathering procedures. This methodology is also useful for studies on diseases or conditions with a low prevalence.

Indirect sampling (nexuses): this is used to analyze social networks according to the relationship or nexus among the different members forming the network.

Calibration of estimators: this permits reducing estimator error, thereby offering superior results precision and adjusting sample weights as a function of relevant population characteristics, ensuring superior estimation validity.

The aims of this doctoral thesis were to review and develop this last methodology, to apply it in population-based health surveys with undercoverage problems and to assess different calibration approaches according to their standard errors. The research question was whether the application of calibration adjustments to estimate the prevalence of a disease increases the validity of its generalization and improves the accuracy of the estimation.

DEVELOPMENT This doctoral thesis is structured into the following four parts:

Part I Research by sampling:

Sampling is one of the research methodologies most widely used by companies, institutions, statistical offices, and public organisms, especially in well-known *population-based surveys*.

This first part deals with health surveys ([chapter 1](#)); it begins by connecting some of the most relevant *population-based health surveys*, continues by describing the basic fields of knowledge that form the *development of a survey*, and ends by centering on the main reasons for the implementation of a survey and enumerating some of the main non-sampling errors that determine its quality. In addition, it reports on the opportunities provided by *mHealth or Big Data* and introduces the concept of *estimator calibration* for the first time.

Part II Indirect estimation methods:

Auxiliary information is that provided by surveys on variables related to the study variable, which can in turn also be obtained for the whole study population. The way in which this auxiliary information is used gives rise to different parameter estimation methods, making them more accurate and valid.

This second part is divided into three chapters. The first ([chapter 2](#)), on auxiliary information, begins with the historical development of the main methods used to include auxiliary information in the estimation stage, continues by explaining when to use it, and ends by introducing the most frequently used *indirect estimators*. The second chapter ([chapter 3](#)) introduces the method of *empirical likelihood* and modifications of its function, (giving rise for example to the model-calibrated pseudo-empirical likelihood estimator) and describes its application in different sample designs. Finally, the third chapter ([chapter 4](#)) introduces *non-parametric regression* and its advantages and estimators, such as the model-assisted or non-parametric calibration. This chapter and this second part end with the utilization of these estimators for both quantitative and qualitative auxiliary information.

Part III Calibration of estimators:

This is one of the most recent techniques for more valid and accurate parameter estimation. It offers a general framework for weighting with a simple and practical approximation to incorporate auxiliary information into the estimation.

This third part describes the state of the art in calibration techniques; it also comprises three chapters. The first ([chapter 5](#)) reports how to construct an estimator of calibration and its conditions, and describes different estimators as a function of the distance selected and the construction process. The second ([chapter 6](#)) develops the case of calibration

applied to *imperfect frames*, i.e., when there are coverage errors. Thus, it describes the different types of errors and how to reduce them for one or multiple imperfect frames, using calibration estimators. Finally, the third chapter ([chapter 7](#)) briefly outlines some of the specific Software packages for the calibration of estimators, focusing on free R Software and its calibration-related packages.

Part iv Application of the calibration methods:

Statistical agencies increasingly apply estimator calibration methodologies for their main surveys.

This last part includes two chapters on the application of calibration in two real population-based surveys. Calibration is applied to estimate the prevalence of *infertility* in the first chapter ([chapter 8](#)) and to estimate the prevalence of *disabling chronic pain* in the second ([chapter 9](#)). Both chapters are structured according to the sections corresponding to an original scientific paper in the Health Sciences field. Accordingly, the first begins by introducing the theoretical framework of sterility and its prevalence and continues with the methodology and analysis of the *Infertility Survey* (Huelva, 2011) and the results obtained, finally analyzing the gain in accuracy after applying linear calibration. The second introduces the theoretical framework of chronic pain, its prevalence, and associated factors and continues with the methodology and analysis of the *Andalusian Health Survey* (2011) and the results obtained, analyzing the gain in accuracy for disabling chronic pain prevalence after applying linear calibration and the assisted model. Both chapters end with a discussion of the results and study limitations and strengths, and conclusions.

The above parts are completed with an [Appendix A](#), describing the code used in statistic R Software as well as a section giving all the references used.

PRESENT AND FUTURE CONTRIBUTIONS This is the first time that a calibration technique has been applied to the surveys used in this work and, to the best of our knowledge, the first time that model-assisted calibration has been applied to a population-based health survey. This approach is being increasingly adopted in social sciences, especially in studies with small samples or with coverage or non-response problems. The methodology developed in this doctoral thesis may serve as a model for application in other epidemiological studies.

Some of the contents gave rise to the presentation of various congress papers and the following *original scientific papers*:

«Calibrated prevalence of infertility in 30- to 49-year-old women according to different approaches: a cross-sectional population-based study”

published in *Human Reproduction* (FI: 4,585; D1; Ranking: 5/79; [A. Cabrera-León et al., 2015](#)),

«Disabling chronic pain: characteristics, multimorbidity, and associated factors based on a face-to-face cross-sectional population study” under review for *Pain Medicine* (FI: 2,34; Q1; Ranking: 37/154; [Andrés Cabrera-León, Cantero-Braojos, et al., 2016](#)) and

«Calibrated prevalence of Disabling chronic pain» under review for *Pain* (FI=5,2; D1; Ranking: 2/30; [Andrés Cabrera-León, Rueda, et al., 2016](#)).

The PhD candidate also presented this doctoral thesis at the *Conference on Public Health Excellence Research* (Menorca, September 2015), where it was very highly rated by the reviewers. They identified promising possibilities for application in epidemiology, not only in observational population-based cross-sectional studies, as in the present doctoral thesis, but also in other populations (e.g., hospitals or schools) and epidemiologic designs (e.g., case-controls or cohorts).

In this respect, the PhD candidate is leading several studies on which calibration technique is also applied. The aim of one of them is to estimate, using calibration, the change in a variable (self-rated general health) measured at two different times in the same cross-sectional survey (Spanish Health Survey) and to compare the results with those obtained using a different statistical technique (exact matching and propensity score-matching). The candidate has also participated in the different editions of the Andalusian Health Survey, adjusting the design weights through calibration in its latest editions (2011 and 2015/16). The aim of another study is to measure the self-rated happiness of healthcare professionals from two different populations (Andalusia and Colombia), applying propensity score-matching and calibration to reduce the selection bias as well as problems of undercoverage and non-response. He is also participating in another initiative called «[The Health of Humanity Project](#)». It researches on health-rated through a database with information from 116 countries.

CONCLUSION *Calibration of estimators* is presented as a powerful statistical methodology for application in the field of *Health Sciences*. With wide statistical and computer development, it comprises multiple techniques that improve the *accuracy and validity of estimations*, providing consistency in the numbers presented and in those from other sources.

RESUMEN

INTRODUCCIÓN La tesis doctoral a defender se titula «*Nuevas técnicas de investigación por muestreo aplicadas a encuestas de salud: calibración de estimadores*».

Uno de los principales requisitos en la *investigación científica* es conseguir la máxima exactitud al medir. La variación en las medidas es propia de la medicina y de cualquier otra ciencia biológica, pero hay otras fuentes de variación que se mezclan con la anterior y que vienen dadas por el propio acto de la medición. Estas son: la *precisión* y la *validez*. Así pues, por un lado la precisión (también conocida como fiabilidad, reproducibilidad, concordancia o acuerdo) está asociada al *error aleatorio*, siendo por tanto impredecible e inevitable, aunque estimable. Por otro lado, la validez está asociada al error sistemático o *sesgo*, y es predecible y corregible. Se suele relacionar la estimación y control de la precisión con la *Estadística*, y las de la validez con la *Epidemiología*.

Algunas de las metodologías de investigación por muestreo desarrolladas en los últimos 25 años que, por un lado, minimizan el sesgo y, por otro, aumentan la precisión de las estimaciones realizadas sobre parámetros medidos a través de muestras, son:

Modelos de respuesta aleatorizada: útiles para aumentar la protección del anonimato, consiguiendo así una mayor fiabilidad y tasa en la respuesta.

Marcos múltiples: permiten superar problemas debidos a la falta de cobertura de determinados marcos muestrales, incorporando distintos procedimientos para la recogida de datos. También esta metodología es útil para realizar estudios sobre enfermedades o condiciones poco prevalentes.

Muestreo indirecto (de nexos): es utilizado para analizar redes sociales mediante la relación o nexo que existe entre los diferentes miembros que la componen.

Calibración de estimadores: permite reducir el error de los estimadores, ofreciendo así mayor precisión en los resultados, y ajustar los pesos muestrales en función de características poblacionales relevantes, asegurando así mayor validez en las estimaciones.

Los *objetivos de investigación* de esta tesis doctoral fueron revisar y desarrollar esta última metodología sobre calibración de estimadores, aplicarla a encuestas poblacionales de salud con problemas de subcobertura y evaluar diferentes aproximaciones de calibración de acuerdo con sus errores estándar. Así pues, se trata de responder a la *pregunta de*

investigación de si los ajustes por calibración aplicados a la estimación de la prevalencia de enfermedades aumentan la validez de su generalización así como su precisión.

DESARROLLO Para ello, esta tesis doctoral se estructura en las siguientes cuatro partes:

Parte I La investigación por muestreo:

El muestreo es una de las metodologías de investigación más ampliamente utilizada por empresas, instituciones, oficinas de estadística y organismos públicos, sobre todo a través de las conocidas encuestas de base poblacional.

Esta primera parte contiene el **Capítulo 1** que comienza enlazando a algunas de las *encuestas poblacionales de salud* más relevantes, continúa describiendo las áreas básicas de conocimiento que componen el *desarrollo de una encuesta* y finaliza deteniéndose en las principales razones que justifican la realización de una encuesta, así como describiendo algunos de los principales errores ajenos al muestreo que determinan su calidad. Además, se mencionan las oportunidades que brindan la llamada *mHealth* o el *Big Data* y se introduce por primera vez el concepto de *calibración de estimadores*.

Parte II Métodos de estimación indirectos:

La *información auxiliar* es la proporcionada por las encuestas sobre variables relacionadas con la variable de estudio que, a su vez, también pueden ser obtenidas para la totalidad de la población de estudio. La forma de utilizar esta información auxiliar da lugar a diferentes métodos de estimación de los parámetros, haciéndolos más precisos y válidos.

Esta segunda parte se divide en 3 capítulos. El **Capítulo 2** comienza con un desarrollo histórico de los primeros métodos utilizados para incorporar la información auxiliar en la fase de estimación, continúa explicando cuándo utilizarla y finaliza introduciendo los *estimadores indirectos* más utilizados. El **Capítulo 3** introduce el método de *verosimilitud empírica*, las modificaciones de su función (que da lugar por ejemplo al estimador de verosimilitud pseudo empírica modelo-calibrado) y su aplicación bajo diferentes diseños muestrales. Por último, el **Capítulo 4** introduce la *regresión no paramétrica*, sus ventajas y estimadores como el de modelo asistido o el de calibración no paramétrico. Este capítulo, así como esta segunda parte, finaliza con la utilización de esos estimadores no solo con información auxiliar cuantitativa, sino también cualitativa.

Parte III Calibración de estimadores:

Se trata de una de las más recientes técnicas para hacer más válida y precisa la estimación de un parámetro. Ofrece un marco general so-

bre ponderación con una *aproximación simple y práctica* para incorporar información auxiliar a la estimación.

Esta tercera parte desarrolla el estado del arte de las técnicas de calibración y se compone igualmente de 3 capítulos. El [Capítulo 5](#) introduce cómo construir un *estimador de calibración*, sus condiciones y diferentes estimadores en función de la distancia elegida así como del proceso de construcción. El [Capítulo 6](#) desarrolla el caso de la calibración aplicada a *marcos imperfectos*, es decir, cuando se dan errores de cobertura. Así pues, se describen los diferentes tipos de esos errores y cómo reducirlos para los casos de uno o múltiples marcos imperfectos a través de los estimadores de calibración. Por último, el [Capítulo 7](#) describe brevemente algunos de los *Softwares* específicos para la técnica de calibración de estimadores, desarrollando en profundidad el Software libre *R* y sus paquetes relacionados con calibración.

Parte iv Aplicación de los métodos de calibración:

Son cada vez más las agencias estadísticas que aplican las metodologías de calibración de estimadores para sus principales encuestas.

Esta última parte consta de dos capítulos correspondientes a la aplicación de la calibración a dos casos reales de encuestas de base poblacional. El [Capítulo 8](#) aplica la calibración en marcos imperfectos para estimar la prevalencia de la *esterilidad* y el [Capítulo 9](#) para estimar la prevalencia del *dolor crónico discapacitante*. Ambos capítulos se estructuran en los apartados comunes a un artículo científico original en el ámbito de Ciencias de la Salud. Así pues, el [Capítulo 8](#) comienza introduciendo el marco teórico de la esterilidad y su prevalencia, continúa con la metodología y análisis de la Encuesta de Esterilidad (Huelva, 2011) para dar paso a los resultados obtenidos y realizando una comparativa en la ganancia de precisión tras aplicar la calibración lineal. El [Capítulo 9](#) comienza introduciendo el marco teórico del dolor crónico, su prevalencia y factores asociados; continúa con la metodología y análisis estadístico de la Encuesta Andaluza de Salud (2011) para dar paso a los resultados obtenidos. También se realiza una comparativa en la ganancia de precisión de la prevalencia del dolor crónico discapacitante según *calibración lineal y modelo asistida*. Ambos capítulos terminan con la discusión de los resultados, las limitaciones y fortalezas, y con las conclusiones.

Las partes anteriores se completan con un [Apéndice A](#) que incluye el *código* utilizado en el Software estadístico *R* así como con un apartado con toda la *bibliografía* utilizada.

CONTRIBUCIONES ACTUALES Y FUTURAS Esta es la primera vez que la técnica de calibración se ha aplicado a las encuestas utilizadas en este trabajo y, hasta dónde sabemos, la primera vez que la calibración

modelo-asistida se ha aplicado a una encuesta poblacional de salud. Esta aproximación está siendo cada vez más adoptada en las ciencias sociales, especialmente en estudios con muestras pequeñas o con problemas de cobertura o de falta de respuesta. La metodología desarrollada en esta tesis doctoral puede servir como modelo para ser aplicado a otros estudios epidemiológicos.

Parte del contenido de esta tesis ha dado lugar a la presentación de varias comunicaciones y a la producción de los *artículos científicos originales*:

«Calibrated prevalence of infertility in 30- to 49-year-old women according to different approaches: a cross-sectional population-based study» publicado en *Human Reproduction* (FI: 4,585; D1; Ranking: 5/79; [A. Cabrera-León y col., 2015](#)),

«Disabling chronic pain: characteristics, multimorbidity, and associated factors based on a face-to-face cross-sectional population study» en revisión por *Pain Medicine* (FI: 2,34; Q1; Ranking: 37/154; [Andrés Cabrera-León, Cantero-Braojos y col., 2016](#)) y

«Calibrated prevalence of Disabling chronic pain» en revisión por *Pain* (FI=5,2; D1; Ranking: 2/30; [Andrés Cabrera-León, Rueda, et al., 2016](#)).

La tesis fue también presentada por el candidato a doctor en el *Encuentro para la Excelencia de la Investigación en Salud Pública* (Menorca, septiembre 2015), con muy buena valoración de los evaluadores. Identificaron prometedoras oportunidades de aplicación a la epidemiología, no solamente en estudios en población general, observacionales y transversales, como en la presente tesis doctoral, sino también en otras poblaciones (p.ej. hospitalarias o escolares) y a otros diseños epidemiológicos (p.ej. casos y controles o cohortes).

En este sentido, además de los trabajos anteriores, el candidato a doctor está al frente de otros estudios en los que la técnica de calibración también es aplicada. Uno de ellos está relacionado con comparar diferentes estimaciones de la varianza en la calibración por dominios. Otro en analizar, a través de la calibración, el cambio de una variable (salud general autopercibida) medida en dos tiempos distintos de la misma encuesta poblacional transversal (Encuesta Nacional de Salud de España) y comparar sus resultados con los obtenidos mediante otra técnica estadística (emparejamiento exacto y mediante propensity score). Relacionado también con encuestas poblacionales, el candidato ha participado en las diferentes ediciones de la Encuesta Andaluza de Salud, calibrando los pesos de diseño en la edición de 2011 y con previsión de hacerlo también en la de 2015/16. Otra investigación analiza la satisfacción autopercibida con la vida a partir de dos poblaciones de profesionales sani-

tarios (Andalucía y Colombia), aplicando inicialmente emparejamiento mediante propensity score y, posteriormente, calibración, con el objetivo de reducir el sesgo de selección así como problemas de cobertura y de no respuesta de las muestras. También está participando en una iniciativa denominada **The Health of Humanity Project**, que analiza la salud autopercebida a partir de la información agregada de 116 países.

CONCLUSIÓN La *calibración de estimadores* se presenta como una metodología estadística poderosa para ser aplicada en el ámbito de las *Ciencias de la Salud*. Con amplio desarrollo estadístico e informático, comprende múltiples técnicas que mejoran la *precisión y validez* de las estimaciones cuando se dan errores ajenos al muestreo, como el de cobertura, proporcionando consistencia entre las cifras presentadas y las procedentes de otras fuentes.

PUBLICACIONES

Algunas ideas, tablas y figuras han aparecido previamente en las siguientes publicaciones:

Arcos, Antonio, José Miguel Contreras y María del Mar Rueda

2014 "A novel calibration in social surveys", *Sociological Methods and Research*, 43, pág. 465-489.

Arcos, Antonio, María del Mar Rueda y David Molina

2013 "Calibration in Complex Survey using R", en *International Conference on USER!*

Cabrera-León, Andrés y Miguel Ángel Cantero-Braojos

2015 "Prevalencia, factores asociados e impacto del dolor crónico limitante: un estudio transversal de base poblacional", en *Congreso de la Sociedad Española de Rehabilitación y Medicina Física*.

Cabrera-León, Andrés, Miguel Ángel Cantero-Braojos, Llenalia Garcia-Fernandez y Juan Antonio Guerra-de-Hoyos

2016 "Disabling chronic pain: characteristics, multimorbidity, and associated factors based on a face-to-face cross-sectional population study", *Pain Medicine* (FI: 2,34; Q1; Ranking: 37/154), en revisión.

Cabrera-León, Andrés, Vicente Lopez-Villaverde, María del Mar Rueda y María Natividad Moya-Garrido

2015a "Calibrated prevalence of infertility in 30- to 49-year-old women according to different approaches: a cross-sectional population-based study", *Human Reproduction* (FI=4.6; D1; Ranking: 5/79), 30, 11, pág. 2677-2685, DOI: [10.1093/humrep/dev226](https://doi.org/10.1093/humrep/dev226), <http://www.ncbi.nlm.nih.gov/pubmed/26370663>.

2015b "Using calibration techniques to analyze infertility in women of childbearing age", en *Sixth International Conference on Europe Survey Research Association*.

Cabrera-León, Andrés, María del Mar Rueda y Miguel Ángel Cantero-Braojos

2016 "Calibrated prevalence of Disabling chronic pain", *Pain* (FI=5,2; D1; Ranking: 2/30), en revisión.

- Iglesias, Diego y María del Mar Rueda
2010 *El método de verosimilitud empírica. La estimación de proporciones y su aplicación a un caso particular*, Trabajo de investigación fin de Máster, Universidad de Granada, pág. 35.
- Lundström, Sixten y Carl Erick Särndal
2001 *Estimation in the Presence of Nonresponse and Frame Imperfections*, Statistics Sweden, pág. 172.
- Martínez Puertas, Sergio
2003 *Extensión del método de calibración a parámetros funcionales (Tesis doctoral)*, Universidad de Almería.
- Martínez, Sergio, Antonio Arcos, Helena Martínez y Sarjinder Singh
2015 "Estimating population proportions by means of calibration estimators", *Revista Colombiana de Estadística*, 38, 1, pág. 267-293.
- Molina, David, Antonio Arcos y María del Mar Rueda
2012 *Estimation in the Presence of Nonresponse and Frame Imperfections*, Trabajo de investigación fin de Máster, Universidad de Granada, pág. 62.
- Pérez, Carlos y María de Blas Portero
2015 "Técnicas de Calibrado dirigidas a encuestas de hogares", en *Escuela de Estadística de las Administraciones Públicas*, Instituto Nacional de Estadística.
- Ranalli, María Giovanna, Antonio Arcos, María del Mar Rueda y Annalisa Teodoro
2015 "Calibration estimation in dual-frame surveys", *Statistical Methods and Applications*, pág. 1-29.
- Rueda, María del Mar e Ismael Sanchez-Borrego
2009 "A predictive estimator of finite population mean using non-parametric regression", *Revista Colombiana de Estadística*, 24, 1, pág. 1-14, DOI: [10.1007/s00180-008-0140-x](https://doi.org/10.1007/s00180-008-0140-x).
- Sanchez-Borrego, Ismael, Jean Didier Opsomer, María del Mar Rueda y Antonio Arcos
2014 "Nonparametric estimation with mixed data types in survey sampling", *Revista Matematica Complutense*, 27, 2, pág. 685-700, DOI: [10.1007/s13163-013-0142-2](https://doi.org/10.1007/s13163-013-0142-2).

*To err is human,
to forgive divine –
but to include errors
in your design is
statistical.*

—Leslie Kish 1977

AGRADECIMIENTOS

Llegar a escribir este apartado ha sido tan difícil como deseado. No imaginé hacerlo y, sin embargo, en ello estoy. Alcanzar esta tesis doctoral, después de tantos años, ha sido posible gracias a muchísimas personas.

En primer lugar, te agradezco a ti, mi Ana, mi compañera, amiga, consejera, amante, amor. Tú me haces seguir soñando y, al mismo tiempo, haces realidad mis mejores sueños, este y todos cuanto haya podido tener, el primero, conocerte. Muchas gracias por tu infinita generosidad en darme tu mejor tiempo para poder dedicarme a este fin, pero sobre todo por tu cariño, confianza, apoyo, consejo. Eres lo mejor, máspreciado y verdadero que tengo. Te amo.

Una vez cogido aire, sigo, cómo no, con vosotras, mis hijas, Ania, Adelina y tú, que siempre me acompañarás. Sois las que me habéis hecho «desconectar» de la tesis en los momentos más necesarios. Siempre sabíais cómo hacerlo y, si no lo conseguíais, volvíais a intentarlo. Veros nacer y crecer es lo más maravilloso que he podido saborear en mi vida. Os amo.

Mis padres, Paqui y Paco, sois el espejo en el que siempre he querido verme. Esta tesis es tan mía como vuestra. Además de los que debemos hacerlo, ¡sois los únicos que os la habéis leído entera! Contar con vuestra experiencia, sinceridad, comprensión y cariño en los momentos más importantes de mi vida ha sido un alivio, una seguridad y, en definitiva, una tremenda suerte, sois la mejor de las loterías.

Y, qué decir de vosotros, mis hermanos, Manuel David y Lucas. Gracias por sacarme de mi rutina doctoral, por hacerme coger distancia de la tesis para retomarla con más ganas. Sois un orgullo para mi, es un honor ser vuestro hermano mayor. Manuela, también te doy las gracias a ti, qué bien tenerte con mi hermano. La ilusión por verte a ti y a tus hijas, mis sobrinas Vega y Julia, compensaron cualquier momento doctoral.

Claro que mis agradecimientos familiares no acaban aquí. Gracias a todos los Leones y algún/a que otro/a Cabrera, especialmente a aquellos que nos dejaron durante el desarrollo de esta tesis; a mis queridos tíos Ángel y Miguel, también mi padrino. A mi abuela Rosa, ¡cuánto me acuerdo de ti! También agradezco a Ana, la madre de mi Ana, por quererme como uno más de sus hijos y, por lo tanto, por querer lo mejor para mi. Y, cómo no, a su nieta y mi sobrina Griselda, la Voz cantante de la familia, y a su hermana, Ana Victoria, madre, Griselda, y padre, Federico.

Y qué decir de mis amigos. Vosotros sabéis que la tesis no ha sido un capricho, ni una necesidad, sino algo que siempre ha estado ahí, sin saber por qué, para qué y, sobre todo, ¿para cuándo? Más allá de las bromas, os doy las gracias a todos, especialmente a Marga, nuestra chachipén gachí fetén, a Sheila, my English Karma, a Migue, mi Karma español, a los Churreteos de playa, amigas desde la infancia, y a los JoaKings, con Bea y Sergio al frente.

Bueno, dejo la parte más emotiva y paso a la más profesional.

Gracias a mis Directores por hacer realidad esta tesis doctoral. A María del Mar, por, efectivamente, ejercer de Directora. Esto que parece tan sencillo, sé, por mi ya madura edad, que no lo es en absoluto. Tu precisa orientación, gran conocimiento y cercano apoyo fue lo que necesité desde el primer momento. Como sabes, he participado como estadístico en muchas tesis doctorales y me siento tremendamente afortunado tenerte como Directora, en este trabajo y en los que ojalá puedan venir. Continuo con Alex, ¡qué decir! Pues sí, muchísimas gracias por haberme acogido en tu vida, por tu apoyo en esta tesis. Como sabes, la estancia en tu centro de innovación e investigación en Toronto fue, profesional y personalmente, una experiencia vital para mi y mi familia. Allí, me desprendí, literal y figuradamente, de las gafas que llevaba desde que empecé la Universidad. Y tú, Antonio, muchas gracias por estar en el momento justo, por ser compañero y jefe al mismo tiempo, y, de manera muy especial, por construir conmigo el escenario más favorable para conseguir finalizar mi tesis.

También quiero agradecer a mis compañeros de la Escuela Andaluza de Salud Pública, sobre todo a aquellos que siguen sin decaer, a pesar de los pesares. A Carlos, exconsejero Delegado y, posteriormente, a Joan Carles, Director, por respetar y aprobar mis decisiones, por insólitas que parecieran. A Cecilia, mi compañera y amiga desde chiquiticos en la EASP, aire fresco, limpio, persona íntegra, profesional excelente, mujer moderna y luchadora. Ya te lo he dicho y ahora te lo escribo, conocerte es lo mejor que ha pasado en la EASP. También quiero agradecer a mis compañeros de OPIMEC, Begoña, Diana, Eva, Ángel y Antonio, ellos saben muy bien lo que deseaba esta tesis; y a Alicia, Amets, Eva,

Julia, María y Pablo por acogerme tan bien en la sala. Igualmente, gracias a Carmen Sigler por ayudarme con el diseño de este y otros tantos trabajos, así como a Inmaculada Aguilera y Carlos Bouza por su valiosa revisión. Por último, a los colegas que han participado en los artículos relacionados con esta tesis (Llenalia García, María Natividad Moya, Miguel Ángel Cantero y Juan Antonio Guerra), al equipo de investigación que conforma el Subprograma de Crisis y Salud del CIBER de Epidemiología y Salud Pública, así como a los numerosos doctorandos y sus directores de tesis con los que he colaborado, por haberme enseñado el contexto doctoral y por convertirlo en un reto personal más que profesional.

Finalizo mis agradecimientos a la música, de manera muy especial a todas las personas que hacen posible Radio Clásica y Radio 3. Destacar a las de los programas de Músicas posibles, Saltamontes y el Vuelo del Fénix, muchas gracias por conseguir hacer más llevaderas las horas más agotadoras del día y de la noche.

Esta sumatoria en mi muestra aleatoria e intencional de la variable calibrada de agradecimientos tiende a infinitos besos y abrazos, resultando en un ¡os deseo lo mejor, siempre!

Granada, febrero de 2016

ÍNDICE GENERAL

I	LA INVESTIGACIÓN POR MUESTREO	1
1	ENCUESTAS POBLACIONALES DE SALUD	3
1.1	Introducción	3
1.2	Desarrollo de una encuesta	6
1.3	Justificación para realizar encuestas	12
1.4	Calidad de las encuestas	14
1.5	Utilidad de las encuestas	19
II	MÉTODOS DE ESTIMACIÓN INDIRECTOS	21
2	CUESTIONES GENERALES	23
2.1	Introducción	23
2.2	Desarrollo histórico	24
2.3	Estimadores indirectos clásicos	27
2.3.1	Definición del estimador general de regresión	27
2.3.2	Estimador de razón o por cociente	29
2.3.3	Estimador de regresión	33
2.3.4	Estimador de diferencia	34
3	VEROSIMILITUD EMPÍRICA	35
3.1	Introducción	35
3.2	El método de verosimilitud empírica	35
3.3	El estimador máximo verosimil empírico	36
3.4	Modificaciones de la función de verosimilitud	36
3.5	Aplicación a diferentes diseños muestrales	38
4	REGRESIÓN NO PARAMÉTRICA	45
4.1	Introducción	45
4.2	Enfoques basados en modelos y modelo-asistido	46
4.2.1	Estimador modelo-asistido	47
4.2.2	Estimador basado en el modelo	47
4.3	Estimador de calibración no paramétrico	48
4.4	Estimador no paramétrico con discontinuidades	48
4.5	Variables auxiliares cuantitativas y cualitativas	49
III	CALIBRACIÓN DE ESTIMADORES	51
5	CUESTIONES GENERALES	53
5.1	Introducción	53
5.2	Construcción de un estimador de calibración	53
5.3	Modificaciones en los estimadores de calibración	57
5.3.1	Estimadores para una familia de distancias	57

5.3.2	Estimadores basados en una forma funcional	60
5.3.3	Estimadores de calibración cosméticos	62
5.3.4	Estimadores de calibración modelo-asistido	65
5.3.5	Estimadores de calibración para la función de distribución y cuantiles	69
5.3.6	Estimadores de calibración en dominios	71
5.4	Otros trabajos sobre calibración	73
5.5	Ventajas y comparación de la calibración	75
6	CALIBRACIÓN PARA MARCOS IMPERFECTOS	77
6.1	Introducción	77
6.2	Un Marco Imperfecto	80
6.3	Varios marcos imperfectos	82
6.3.1	Estimación en encuestas con dos marcos	83
6.3.2	Método de calibración dual-frame	86
6.3.3	Método de calibración single-frame	87
7	DESARROLLO INFORMÁTICO PARA CALIBRACIÓN	89
7.1	Introducción	89
7.2	Paquetes estadísticos relacionados con calibración	89
7.3	Calibración en R	92
IV	APPLICATION OF CALIBRATION METHODS	101
8	CALIBRATED PREVALENCE OF INFERTILITY	103
8.1	Introduction	103
8.2	Methods	104
8.2.1	Design	104
8.2.2	Statistical analysis	107
8.3	Results	111
8.4	Discussion	113
9	CALIBRATED PREVALENCE OF DISABLING CHRONIC PAIN	121
9.1	Introduction	121
9.2	Methods	123
9.2.1	Design	123
9.2.2	Statistical analysis	123
9.3	Results	127
9.4	Discussion	140
V	APÉNDICE	151
A	CÓDIGO INFORMÁTICO EN R	153
A.1	Prevalencia calibrada de esterilidad	153
A.2	Prevalencia calibrada de dolor crónico discapacitante	154
	BIBLIOGRAFÍA	173

ÍNDICE DE FIGURAS

Figura 1.1	Proceso de desarrollo de una encuesta	7
Figura 1.2	Simulación de la falta de respuesta	17
Figura 2.1	Métodos de estimación indirectos: enfoques	25
Figura 6.1	Error de sobrecobertura	78
Figura 6.2	Error de subcobertura	79
Figura 6.3	Contexto de trabajo en marcos duales.	83
Figure 8.1	Infertility approaches	108
Figure 8.2	Conditions for the infertility calibration variables	111
Figure 8.3	Prevalence of infertility	112
Figure 8.4	Reductions in variance and prevalence	115
Figure 9.1	Conditions for the DCP calibration variables	126
Figure 9.2	Prevalence of chronic diseases	135
Figure 9.3	Multimorbidity and chronic pain	137
Figura 9.4	Disabling chronic pain associated factors	146
Figura 9.5	Non-disabling chronic pain associated factors	147
Figura 9.6	Observed biases (calibration variables)	148
Figura 9.7	Calibrated prevalence of disabling chronic pain	148
Figura 9.8	Percentage reductions after calibration	149

ÍNDICE DE TABLAS

Tabla 1.1	Encuestas poblacionales de salud y métodos de estimación indirectos	5
Tabla 1.2	Programas informáticos para encuestas	12
Tabla 5.1	Ejemplos de funciones $G_k(\omega_k, d_k)$.	59
Table 8.1	Sample distribution	106
Table 8.2	Infertility approaches	109
Table 8.3	Selection of infertility variables to calibrate	110
Table 8.4	Agreement for infertility variables	110
Table 8.5	Socio-demographic features	114
Table 9.1	Study variables	124
Table 9.2	Characteristics of the study population	128
Table 9.3	Life habits	129
Table 9.4	Disabling chronic pain	131
Table 9.5	Factors associated with chronic pain	133

Table 9.6	Disabling chronic pain and chronicity	134
Table 9.7	Observed biases of the calibration variables	139
Table 9.8	Calibrated prevalence of disabling chronic pain	140

LISTADO DE CÓDIGOS

Listing 7.1	Función kottcalibrate	94
Listing 7.2	Función CalDF	94
Listing 7.3	Función CalSF	94
Listing 7.4	Función JackCalDF	94
Listing 7.5	Función JackCalSF	94
Listing 7.6	Función WeightsCalDF	97
Listing 7.7	Función WeightsCalSF	97
Listing 7.8	Función calibWeights	97
Listing 7.9	Función calib	97
Listing 7.10	Función calibev	97
Listing 7.11	Función checkcalibration	97
Listing 7.12	Función gencalib	97
Listing 7.13	Función calibrate	99
Listing A.1	Infertility: starting R	155
Listing A.2	Estimation without calibrating	156
Listing A.3	Linear calibration 1	157
Listing A.4	Linear calibration 2	158
Listing A.5	Linear calibration 3	159
Listing A.6	Linear calibration 4	160
Listing A.7	Linear calibration 5	161
Listing A.8	Calibration estimations	162
Listing A.9	Reductions after calibration	163
Listing A.10	Disabling Chronic Pain: starting R	164
Listing A.11	Estimation without calibrating	164
Listing A.12	Linear calibration (marginals)	165
Listing A.13	Linear calibration (crossings)	166
Listing A.14	Raking calibration	167
Listing A.15	Model-assisted calibration	168
Listing A.16	Model-assisted calibration (raking)	169
Listing A.17	Calibration estimations	170
Listing A.18	Reductions after calibration	171

ACRÓNIMOS

ADL	Activities of Daily Living
BRR	Balanced Repeated Replication
CAPI	Computer-Assisted Personal Interview
CC	Chronic Cephalalgia
CI	Confidence Interval
CP	Chronic Pain
CV	Coefficient of Variation
DCP	Disabling Chronic Pain
EAS	Encuesta Andaluza de Salud
EASP	Escuela Andaluza de Salud Pública
ECM	error cuadrático medio
ECV	Encuesta de Condiciones de Vida
EDAD	Encuesta de Discapacidad, Autonomía Personal y Situaciones de Dependencia
EDADES	Encuesta Domiciliaria sobre Alcohol y Drogas en España
ENSE	Encuesta Nacional de Salud de España
ESHRE	European Society for Human Reproduction and Embryology
ESTUDES	Encuesta Estatal sobre Uso de Drogas en Enseñanza Secundaria
EPA	Encuesta de Población Activa
ESHRE	Sociedad Europea para la Reproducción Humana y Embriología
GREG	regresión generalizada
INE	Instituto Nacional de Estadística
MAS	muestreo aleatorio simple

MSP Musculoskeletal Pain

nDCP non-Disabling Chronic Pain

OR Odds Ratio

PEL pseudo-verosimilitud empírica

PML pseudo-verosimilitud máxima

SAS Statistical Analysis System

SPSS Statistical Package for the Social Sciences

VE verosimilitud empírica

WHO World Health Organization

Parte I

LA INVESTIGACIÓN POR MUESTREO

El muestreo es una de las metodologías de investigación más ampliamente utilizada por empresas, instituciones, oficinas de estadística y organismos públicos, sobre todo a través de las conocidas encuestas de base poblacional. Esta primera parte contiene el [Capítulo 1](#) que comienza enlazando a algunas de las *encuestas poblacionales de salud* más relevantes, continúa describiendo las áreas básicas de conocimiento que componen el *desarrollo de una encuesta* y finaliza deteniéndose en las principales razones que justifican la realización de una encuesta, así como describiendo algunos de los principales errores ajenos al muestreo que determinan su calidad. Además, se mencionan las oportunidades que brindan la llamada *mHealth* o el *Big Data* y se introduce por primera vez el concepto de *calibración de estimadores*.

ENCUESTAS POBLACIONALES DE SALUD

1.1 INTRODUCCIÓN

Este capítulo describe las características fundamentales de las encuestas poblacionales de salud, especialmente de aquellas con diseño epidemiológico observacional transversal.

Algunas de las encuestas poblacionales más relevantes, desarrolladas en España en el ámbito de la Salud, son las diferentes ediciones de la Encuesta Nacional de Salud de España (ENSE) que desde 1987 viene realizándose (Ministerio de Sanidad, 2013) o las encuestas de salud que cada Comunidad Autónoma aporta a sus respectivos sistemas sanitarios (Ministerio de Sanidad, 2015), como por ejemplo la Encuesta Andaluza de Salud (EAS), con cinco ediciones desde 1999 (Sánchez Cruz y col., 2013). Otras encuestas poblacionales relevantes en España son: la Encuesta de Redes Familiares en Andalucía (Fernández Cordón y Tobío, 2007) realizada en 2005, de ámbito andaluz y con muestras representativas para sus provincias, recoge información sobre características sociodemográficas de las redes familiares, las relaciones que se establecen en las mismas, su tamaño y la necesidad de ayudas y cuidados; la Encuesta de Discapacidad, Autonomía Personal y Situaciones de Dependencia (EDAD) (ediciones de 1999 y 2008) (INE, 2009) realizada en 2008; la Encuesta Domiciliaria sobre Alcohol y Drogas en España (EDADES) (Observatorio Español de Drogas, 2013), encuesta bianual realizada en hogares a población entre 15 y 64 años y de ámbito español; o la Encuesta Estatal sobre Uso de Drogas en Enseñanza Secundaria (ESTUDES), encuesta en centros educativos que viene realizándose bianualmente en España desde 1994 a jóvenes entre 14 y 18 años. En el ámbito internacional son destacables la Encuesta Poblacional de Salud de Canadá (Yergens y col., 2014); la Encuesta Longitudinal de Salud de las Mujeres en Australia (Dobson y col., 2015); la Encuesta Panel de Hogares Británicos (Yu y col., 2015); la Encuesta Longitudinal sobre la Población Infantil en Estados Unidos (Altman y col., 2014) (interrumpida en diciembre de 2014); la Encuesta Europea de Salud (Eurostat, 2013b), que se realiza cada 5 años desde 2006 recogiendo información de personas mayores de 15 años en 17 países europeos; la de Alfabetización de la Salud (Sørensen, Pelikan y col., 2015; Sørensen, Van den Broucke y col., 2013); la Encuesta Social Europea (Eurostat, 2013a) que recoge información cada 2 años desde 2001; la Encuesta de Condiciones de Vi-

da (ECV) (Arora y col., 2015; Goedeme, 2013) con diseño longitudinal que recoge información en más de 30 países europeos sobre ingresos, pobreza, exclusión social y otras condiciones de vida; y la Encuesta Europea de Salud, Envejecimiento y Jubilación (Boersch-Supan y col., 2013; Lumsdaine y Exterkate, 2013), que analiza en profundidad el proceso de envejecimiento poblacional a partir de la información de personas mayores de 50 años de 20 países europeos.

La Tabla 1.1 muestra los métodos de estimación indirectos aplicados por algunas de las anteriores encuestas, es decir, las técnicas que utilizan para incorporar la información auxiliar a las estimaciones de los parámetros de estudio. Así pues, se puede observar que las agencias estadísticas son los organismos que fundamentalmente aplican la metodología de calibración de estimadores, utilizando en su mayoría la distancia lineal o chi-cuadrado y como variables principales de calibración las de sexo, edad y región, aunque depende mucho de la calidad y desagregación de los registros administrativos disponibles en cada país.

Las agencias de estadísticas son las que están implementando en mayor medida las técnicas de calibración de estimadores en encuestas poblacionales, lo cual indica su importancia y a la vez la gran oportunidad de extenderlas a encuestas realizadas por otras organizaciones. Algunos ejemplos son:

- Instituto Nacional de Estadística (INE): además de las ya mencionadas ENSE y ECV, el INE utiliza técnicas de calibración en la Encuesta de Población Activa (EPA) según grupos de edad y sexo, utilizando las proyecciones de población relativas a cada CCAA y referidas a cada trimestre. Están trabajando en estandarizar las variables auxiliares para ofrecer resultados consistentes entre sus encuestas.
- Eurostat: utiliza calibración en la gran mayoría de las encuestas poblacionales que lleva a cabo (como las encuestas SHARE, EU-SILC o EHIS mencionadas anteriormente), utilizando como variables al menos sexo, edad y, dependiendo de cada país, la región.
- Statistics New Zealand (SNZ): con el objetivo de aumentar la calidad, la coherencia y la comparabilidad de los datos de las encuestas básicas y satisfacer las necesidades de los usuarios de manera más eficiente, centran el método de calibración en reunir las principales variables demográficas tales como edad, sexo, origen étnico y que éstas sean coherentes a través de encuestas sociales. La técnica de calibración es utilizada como parte de la estrategia de la Encuesta de Hogares para producir estimaciones de la población, proporcionar información para estimar los puntos de referencia y proporcionar submuestras de las poblaciones de interés.

Tabla 1.1: Encuestas poblacionales de salud y métodos de estimación indirectos

Encuesta poblacional	Método de estimación indirecto
Encuesta Nacional de Salud de España (ENSE) (Ministerio de Sanidad, Servicios Sociales e Igualdad. Instituto Nacional de Estadística)	Calibración lineal según sexo, edad y CCAA. Coeficiente de corrección de falta de respuesta diferencial.
Encuesta Andaluza de Redes Familiares (Instituto de Estadística y Cartografía de Andalucía)	Calibración raking por sexo y grupos quinquenales.
Encuesta Social Europea (ESS) (European Research Infrastructure Consortium)	Postestratificación según grupos edad, sexo, educación y región.
Encuesta sobre ingresos y condiciones de vida (EU-SILC) (Eurostat)	Calibración logit truncado (LO=0,1; UP=10). En España las variables de calibración son sexo, grupos de edad (0-15 y quinquenales en adultos) y tamaño del hogar (excepto Ceuta y Melilla). En otros países depende de la información auxiliar de calidad disponible.
Encuesta sobre alcohol y drogas en España (EDADES) (Ministerio de Sanidad, Servicios Sociales e Igualdad)	Factores simples de ponderación por comunidad autónoma, tamaño de municipio (7 grupos), edad (7 grupos) y sexo (1260 coeficientes de ponderación diferentes).
Encuesta sobre uso de drogas en Enseñanzas Secundarias en España (ESTUDES) (Ministerio de Sanidad, Servicios Sociales e Igualdad)	Factores simples de ponderación por comunidad autónoma, titularidad del centro (público o privado) y tipo de estudios (ESO, Bachillerato, Ciclos Formativos de Grado Medio de Formación Profesional)
Encuesta de Salud, Envejecimiento y Jubilación en Europa (SHARE) (Comisión Europea)	Calibración para hogares e individuos según sexo y grupos de edad (4) cuando es posible según el país correspondiente, y región en 2 países.
Encuesta Europea de Salud (EHIS) (Eurostat).	Calibración según las posibilidades de cada país, al menos por sexo y grupos de edad en decenios, recomendando grupos quinquenales y regiones.

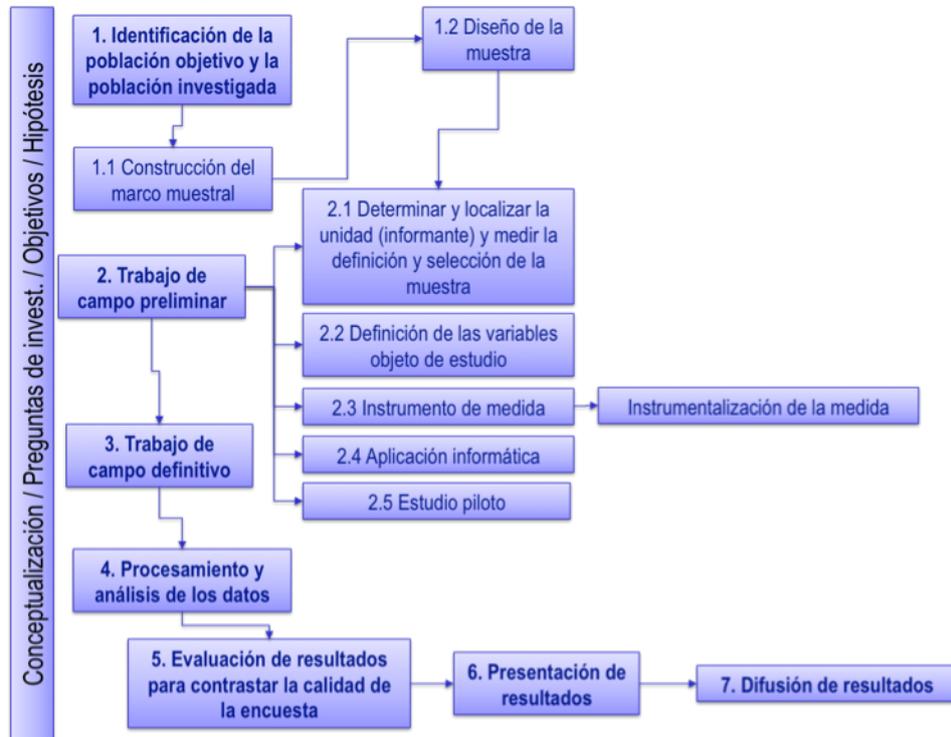
- Office for National Statistics (UK): utiliza la calibración en encuestas sociales tales como la Encuesta de Población Activa, Encuesta de Gastos y Alimentación, la Encuesta General de Hogares y la Encuesta colectiva.
- Statistics Finland: utilizan las técnicas de calibración en la mayoría de estudios sobre población, por ejemplo en la Encuesta de Población Activa o en la Encuesta sobre la Educación de Adultos.
- Statistics South Africa: utilizan calibración en la Encuesta de Población Activa.
- French National Institute of Statistics and Economic Studies (INSEE): utiliza calibración en encuestas sobre los modos de vida y presupuestos.
- Statistique Suisse: utiliza calibración en el Censo Federal de Población y en la Encuesta sobre la Renta y el Consumo.
- **Istat** (Istituto nazionale di statistica, Italia): aplica calibración a la Encuesta de Población Activa y a la Encuesta de Pequeñas y Medianas Empresas, aunque también a otras 20 encuestas más. Para ello, el Istat ha desarrollado un nuevo software en R (RGenesees, ver [Capítulo 7](#)) que está siendo incorporado a otras agencias de estadística, como la escocesa, y evaluado muy favorablemente por las agencias del Reino Unido o de Canadá.

1.2 DESARROLLO DE UNA ENCUESTA

Además de los métodos utilizados para la recogida de la información, son otras muchas las áreas de conocimiento que pueden intervenir en el desarrollo de una encuesta ([Lynn y col., 2005](#)): diseño muestral y de la encuesta, trabajo de campo y otros temas operativos, instrumentos de medida, falta de respuesta y desgaste de la muestra, imputación de datos, errores de medición, ponderación, procesamiento y depuración de la información, análisis estadístico, uso de información auxiliar, ética (como la confidencialidad, el anonimato y el consentimiento) y buenas prácticas para el aseguramiento de la calidad de la encuesta ([Eurostat, 2011](#)). Además, todas estas áreas proporcionan una gran cantidad de materiales y metodologías que posibilitan la adaptación de las encuestas a las características de la población de estudio, a la naturaleza de la información a recoger y a sus posibilidades presupuestarias.

La **Figura 1.1** muestra esquemáticamente las fases más destacadas de la metodología estadística en la investigación por muestreo para el desarrollo de una encuesta poblacional.

Figura 1.1: Proceso de desarrollo de una encuesta



Antes de afrontar la ejecución de cualquier encuesta es fundamental conceptualizarla en función de la temática sobre la que va a recoger información, con el fin de ser capaz de contestar adecuadamente a las preguntas de investigación formuladas, alcanzar sus objetivos y confirmar o descartar las hipótesis planteadas. Cómo de exhaustivo se haga este primer paso, marcará el éxito de la encuesta. Por ejemplo, para alcanzar los objetivos propuestos, se deberá:

- recopilar la información necesaria para cumplirlos,
- definir el motivo de la encuesta,
- identificar la información disponible y aquella complementaria útil para mejorar los procesos de estratificación, selección o estimación,
- valorar los medios materiales o personales,
- establecer los límites presupuestarios y temporales,
- considerar la legislación vigente relacionada y las posibles restricciones administrativas y
- analizar la oportunidad de fechas para su desarrollo.

Una vez enmarcada conceptualmente la encuesta, el primer paso es identificar la población objetivo y sus posibles diferencias con respecto a la investigada debidas, por ejemplo, a subpoblaciones omitidas por no poder acceder a ellas o por falta de respuesta. En este caso se deberá recopilar información adicional que permita disminuir la brecha entre ambas poblaciones. Este primer paso se basa en la construcción del marco muestral, normalmente a partir de infraestructuras estadísticas previas, tratando de depurarlo y completarlo al máximo con el fin de disminuir errores de cobertura. En función de este marco muestral se realiza el diseño de la muestra, donde se deberá describir:

- las unidades de muestreo,
- información complementaria,
- el tamaño de la muestra,
- el tipo de muestreo, selección de la muestra,
- las fórmulas para los estimadores a utilizar y para la estimación de los errores de muestreo,
- el tratamiento de la falta de respuesta y
- el control de otros errores ajenos al muestreo.

En el trabajo de campo preliminar, para instrumentalizar la medida se deberá:

- Formalizar presupuestos así como prever su ejecución y control,
- determinar el método idóneo para la recogida de datos,
- elaborar manuales de instrucción para la cumplimentación de los cuestionarios,
- diseñar e imprimir (si procede) el material de trabajo,
- diseñar la propaganda y la difusión de la encuesta,
- diseñar las múltiples visitas a las unidades informantes para la correcta elaboración de la encuesta,
- preparar cuadros y tablas referentes a la información a recoger y
- seleccionar y entrenar agentes entrevistadores y supervisores.

Con respecto a los métodos de recogida de la información en las encuestas, destacar la evolución que han experimentado con el progreso tecnológico de los últimos años, yendo desde el clásico papel (paper-and-pencil interview, PAPI) hasta el ordenador (computer-assisted personal interviewing, CAPI; computer-assisted telephone interviewing, CATI; computer-assisted self interviewing, CASI), Internet (computer-assisted web interviewing, CAWI) y, los más recientes, los dispositivos y aplicaciones móviles (Pfeffermann y Rao, 2009). Así pues, las aplicaciones informáticas permiten:

- controlar el estado de colaboración del informante,
- asignar el trabajo para cada agente entrevistador,
- obtener indicadores ligados a las recogida de datos,
- actualizar las bases de datos de las unidades muestrales,
- incluir variables testigo que permitan detectar dónde hay que concentrar más los esfuerzos en la última fase de la recogida,
- grabar, depurar y analizar automáticamente los datos y
- conectar los datos proporcionados con los de otras fuentes de información.

Esta última potencialidad de las herramientas informáticas aplicadas a dispositivos móviles, plantea nuevos retos para los países desarrollados y nuevas oportunidades para los países en desarrollo. Así pues, convierten en bidireccional el traspaso de la información entre el informante (entrevistado) y el informado (persona u organismo que recibe la información). Esto es lo que se viene conociendo como *mHealth38* (Franz-Vasdeki y col., 2014; WHO, 2015a) y que se traduce en beneficios directos para las poblacionales de las que se recoge información en términos de:

- incrementar su acceso a la atención sanitaria e información relacionada con la salud, particularmente para poblaciones de difícil acceso,
- mejorar la capacidad para diagnosticar y seguir el desarrollo de enfermedades,
- hacer más oportuna y procesable la información sobre salud y
- expandir el acceso a la educación médica puesta en marcha y el entrenamiento para profesionales de la salud.

Por otra parte, rompiendo con la visión tradicional de recogida y tratamiento de la información, irrumpen en nuestras vidas cada vez con más fuerza las redes digitales, conteniendo ingentes registros detallados de nuestras experiencias diarias que, conjuntamente, comprenden lo que se conoce como *Big Data* (Pentland y col., 2013): dónde vivimos y trabajamos, nivel de actividad, patrones de desplazamientos, hábitos de consumo, qué comemos y bebemos y con qué personas nos relacionamos. Esa información unida a registros de salud, datos demográficos y genéticos, asegura novedosas oportunidades para modelos de salud poblacional no cubiertos, predecir enfermedades de larga duración e identificar aspectos no tradicionales de las intervenciones.

Siguiendo con el desarrollo de una encuesta, el estudio piloto permitirá:

- ensayar el cuestionario en condiciones reales,
- poner a prueba los aspectos fundamentales de la encuesta,
- contrastar la idoneidad del marco,
- resaltar la variabilidad de determinados caracteres,
- intuir la tasa esperada de falta de respuestas,
- comprobar la idoneidad del método de recogida de datos y
- aportar datos sobre el probable coste y duración de la encuesta principal.

En el siguiente paso sobre trabajo de campo definitivo se deberá:

- procesar y depurar automáticamente el cuestionario,
- implantar la información faltante,
- ajustar la no respuesta,
- calcular las estimaciones y sus errores,
- tabular los datos y
- analizar los resultados del trabajo de campo.

Una vez procesados y analizados los datos, se deberá analizar la calidad de la encuesta. Para ello, hay que contrastar las discrepancias entre diseño teórico y aplicado, evaluar los errores ajenos al muestreo y debidos al muestreo, analizar los costes ocasionados, comparar los resultados con otros posibles diseños y contrastar los resultados con otras

encuestas similares, con series históricas, para las diversas segmentaciones territoriales, sectoriales. Los dos últimos pasos están relacionados con la presentación y difusión de resultados. Estos podrán realizarse en soporte papel de modo resumido (Informe ejecutivo), en soporte digital con el grueso de la información de la encuesta, en Internet o en aplicación móvil. Además, podrán publicarse avances previos a los resultados finales, tratando de realizar, en todo momento, una difusión a medida del receptor de la información.

Esta tesis doctoral se centra en dos de los elementos de la anterior [Figura 1.1](#), fundamentales para llevar a cabo la *técnica de calibración de estimadores*. Estos son:

- el diseño de la muestra, que definirá cómo seleccionar la muestra, y
- el análisis de los datos, que ofrecerá estimaciones de parámetros relacionados con las variables de estudio.

En ese sentido se definen los *pesos muestrales* (o factores de elevación) como el número de elementos poblacionales representados por cada elemento de la muestra con respecto a la variable de estudio. La obtención de los pesos en encuestas poblacionales se realiza habitualmente mediante el siguiente proceso:

1. Pesos de diseño.
2. Corrección de falta de respuesta.
3. Reponderación con fuentes externas.

En estos dos últimos puntos es donde la técnica de calibración tiene un papel fundamental.

Por último, la [Tabla 1.2](#) muestra programas informáticos que incorporan módulos relacionados con el desarrollo de encuestas.

Tabla 1.2: Programas informáticos relacionados con encuestas.

MÓDULO	PROGRAMA
Cálculo del tamaño muestral	EpiInfo, SAMPLEPOWER, GRAMMO, SOTAM, R
Cálculo de la potencia estadística	SAMPLEPOWER (SPSS), R
Selección de muestras	SAS, SIGESMU, R
Estimación de errores muestrales	SUDAAN, STATA, CSAMPLE (EpiInfo), PCCARP, R
Análisis estadístico de muestras aleatorias simples	SPSS, SYSTAT, STATISTICA, STATGRAPHICS, BMDP, S-PLUS, SAS, MINITAB, GLIM, STATA, R
Análisis estadístico de muestras complejas	SUDAAN, SPSS, PCCARP, R
Elaboración de Encuestas Web	LIMESURVEY, WUFOO

Como se puede observar, son muchos los programas informáticos que abordan las diferentes fases de una encuesta, siendo el programa informático R ([R Core Team, 2015](#)) el que da respuesta a gran parte de ellas, siendo además gratuito, de ahí que lo hayamos elegido para desarrollar la parte aplicada de esta tesis (más información en el [Capítulo 7](#)).

1.3 JUSTIFICACIÓN PARA REALIZAR ENCUESTAS

Las encuestas proporcionan fuentes de información muy relevantes para la investigación científica y la toma de decisiones en ámbitos como el sanitario, económico, laboral, social o educativo. Algunas de las razones que justificarían la aplicación de una encuesta serían ([Azorín y J. Sánchez-Crespo, 1995](#)):

- La población es tan grande que estudiar todas las unidades excede de las posibilidades de la investigación.
- Uniformidad de la población para obtener buena representatividad.
- Dificultad para recoger la información o para superar la complejidad que supone su uso.
- Reduce el esfuerzo, tanto absoluto como relativo en términos de coste y tiempo (criterio de economía).

- La acuracidad es muy buena (criterio de calidad).
- La formación del personal y la intensidad de los controles y supervisión son altos.
- Constituye la solución de mayor eficiencia en el sentido de coste-beneficio.

Por otra parte, las características deseables para una encuesta serían:

- Acuracidad: proximidad de la estimación al valor verdadero.
- Pertinencia: completar lagunas de conocimiento sobre un fenómeno.
- Oportunidad: puntualidad, rapidez y actualidad.
- Accesibilidad: protección de la intimidad, secreto estadístico, Ley de la Función Estadística Pública.
- Detalle y cobertura: información complementaria a la de una investigación exhaustiva (censos).
- Economía: reducir tiempo y presupuesto del estudio.
- Integración (concepción global de la información) y comparabilidad (con otras informaciones).
- Muestra probabilística (aleatoriedad).

La imposibilidad o desconocimiento para cumplir este último punto en una encuesta ha dado lugar a una gran variedad de tipos de muestreo no probabilísticos como, por ejemplo:

- Sin norma: la muestra se elige sin reflexión ni previsión alguna, de cualquier manera.
- Aplicando criterio: la muestra se elige mediante el criterio razonado de autoridades en la materia de estudio.
- Consecutivo: selecciona a las unidades según van acudiendo (a la consulta, farmacia, banco, etc. . .).
- Por cuotas: la muestra contiene el mismo porcentaje de individuos con ciertas características prefijadas que la población total.
- Semiprobabilístico (superior y/o inferior): la selección de la muestra es probabilística hasta cierto punto del proceso de selección.

A pesar de lo que se pueda pensar, es tal la utilidad de las encuestas que incluso pueden complementar o sustituir fuentes tradicionales de información exhaustivas como los censos. Esto es debido fundamentalmente a que, además de reducir el costo total de la obtención de datos, las encuestas pueden producir mayor precisión adoptando medidas que mejoren la calidad del dato primario. Es por ello que existe una tendencia al uso conjunto de censos y de encuestas como, por ejemplo, en el caso del Censo de Población y Viviendas de España de 2011. A diferencia de los censos anteriores, en esa ocasión se recogió información de una muestra de aproximadamente el 12 % de las viviendas (unos 5,7 millones de personas, 3 millones de viviendas), en todo el territorio español. En cada municipio se extrajo una muestra de tamaño variable, siendo del 100 % en los municipios menores de 200 habitantes y hasta el 9 % en los municipios de más de 1.000.000 de habitantes. La muestra se extrajo a partir de un «fichero precensal» de personas, con información padronal para las distintas personas así como de otros registros administrativos. Esta combinación entre censos basados en registros y grandes encuestas por muestreo también se está aplicando en otros países como Alemania, Holanda, Polonia, Suiza e Israel (más información en [INE 2011](#)).

1.4 CALIDAD DE LAS ENCUESTAS

La credibilidad y significabilidad de los datos que aporta la encuesta, en el análisis de la realidad social, están supeditadas al rigor que se haya puesto en su diseño, ejecución, análisis estadístico e interpretación de la información recogida. Así pues, la calidad de una encuesta ([Biemer y Lyberg, 2003](#); [Cea-D'Ancona, 2005](#)) no debe medirse solo a través del *error muestral*, producido al observar una parte de la población (muestra), sino que el análisis de la precisión de las estimaciones muestrales también exige considerar los *errores ajenos al muestreo*, es decir, aquellos que producen sesgos en las estimaciones. Así pues, mientras los errores de muestreo se pueden medir y disminuyen conforme aumenta el tamaño muestral, los ajenos al muestreo son difíciles de medir y frecuentemente aumentan con la ampliación del tamaño muestral.

Algunos ejemplos de errores ajenos al muestreo son:

- De observación:

SOBRECOBERTURA: qué elementos en la selección de la muestra no pertenecen a la población de estudio.

MEDIDA: instrumentos, pruebas, escalas, ítems, respuestas no adecuados (no fiables, desorganizados o incorrectamente presen-

tados); fuentes secundarias de información no fiables; características personales del entrevistador y su actuación durante la entrevista; estado de ánimo del entrevistado, cansancio, interés, la imagen que quiera proyectar de sí mismo, sus conocimientos del tema, la memoria.

TRATAMIENTO DE LA INFORMACIÓN: errores de grabación, codificación, análisis.

PROCESO: recogida intencional de la información, con fases o procedimientos que la dificultan o condicionan, desarrollo inadecuado de la entrevista (otras personas presentes, interrupciones).

■ De no observación:

SUBCOBERTURA: qué elementos de la población de estudio se han excluido en la selección de la muestra. Tanto el error de sobrecobertura como el de subcobertura, que son consecuencia de marcos imperfectos, se desarrollan específicamente en el [Capítulo 6](#).

FALTA DE RESPUESTA: es uno de los más importantes, teniendo consecuencias directas en ([Groves y col., 2001](#); [RJ Little y Rubin, 2016](#); [Särndal y Lundström, 2005](#)):

- la pérdida de representatividad de la muestra, produciendo mayores sesgos en las estimaciones. Efectivamente:

$$\text{Sesgo} = B(Y) = E\left(\hat{Y}_1\right) - \bar{Y} = \bar{Y}_1 - \bar{Y} = (\bar{Y}_1 - \bar{Y}_2) \frac{N_2}{N}, \quad (1.1)$$

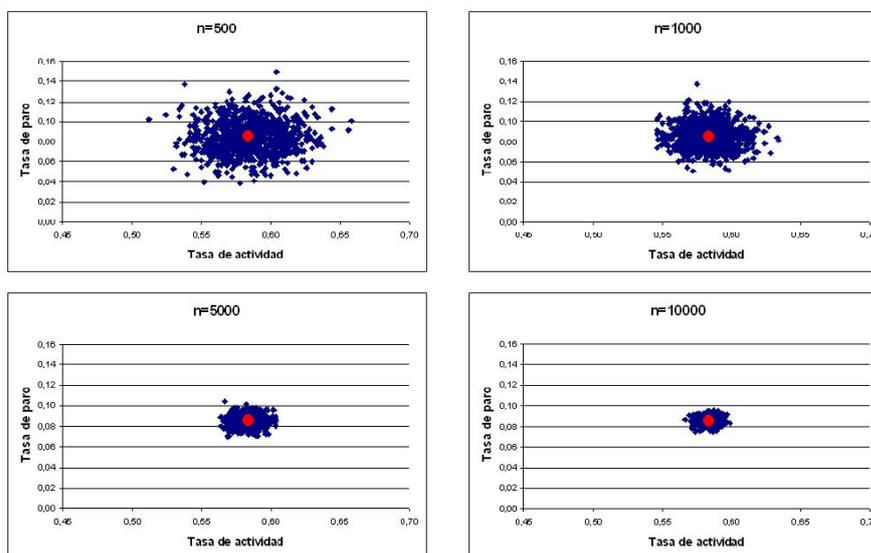
donde Y es la variable de estudio, Y_1 es la población que responde (con N_1 elementos) y Y_2 la que no responde (con N_2 elementos, siendo $N_1 + N_2 = N$). Así pues, el sesgo depende de la diferencia entre la media de la población que responde y la de la población que no responde así como del número de unidades que no responden.

- el aumento de la varianza de las estimaciones, al disminuir el tamaño muestral, y
- el incremento del coste del proyecto, al ser necesario un aumento del tamaño de la muestra en previsión de la pérdida de información.

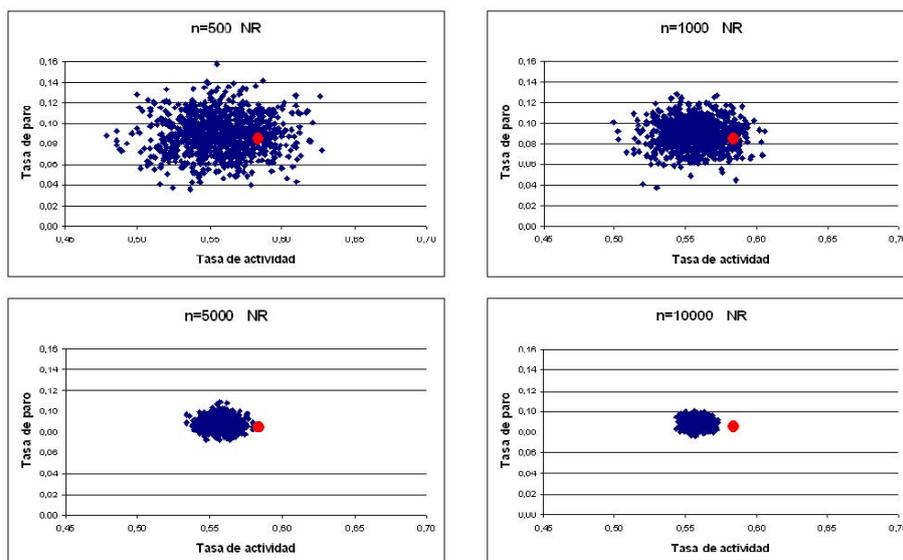
La [Figura 1.2](#) muestra dos simulaciones realizadas por el [INE](#), para mostrar el efecto de la *falta de respuesta*. Así pues, en la [Figura 1.3a](#) se observa cómo las estimaciones de las tasas de paro y actividad de las

muestras (puntos azules) se distribuyen aleatoriamente alrededor de las de toda la población de Lorca (punto rojo), acercándose (es decir, disminuyendo el sesgo) a medida que aumenta el tamaño de muestra. Por el contrario, ante la falta de respuesta ([Figura 1.3b](#)), se observa cómo las estimaciones de las tasas de paro y actividad de las muestras (puntos azules) se distribuyen a la izquierda de las de toda la población de Lorca (punto rojo), habiendo más muestras que se alejan (es decir, aumentando el sesgo) a medida que aumenta el tamaño de muestra.

Figura 1.2: Lorca 2001. Simulación (muestreo aleatorio simple, 1000 iteraciones). Fuente: Instituto Nacional de Estadística.



(a) Tasa de respuesta=100 %



(b) No respuesta: $PR(\text{Varones})=0,7$

En definitiva, los errores ajenos al muestreo pueden afectar gravemente a las posibilidades de inferencia de los datos de la encuesta, siendo el error muestral, en definitiva, insuficiente en su cuantificación.

Algunas actuaciones más consensuadas para mitigar los efectos de estos errores pueden ser (Cea-D'Ancona, 2005):

- Errores de muestreo:

- Incrementar el tamaño de la muestra.
- Selección plenamente aleatoria que garantice la heterogeneidad de la muestra.
- Errores de observación:
 - Errores por el método de encuesta:
 - Elección razonada del método de encuesta, adecuándose al tema de estudio, su complejidad, el contexto social, la población de interés, el tiempo y los recursos disponibles.
 - Mejoras en el diseño del cuestionario.
 - Mejoras en su aplicación.
 - Errores por el cuestionario:
 - Ajustar la formulación de la pregunta y su ubicación al tema y a la población de estudio.
 - Validar o al menos comprobar el correcto funcionamiento del cuestionario mediante un trabajo de campo preliminar o estudio piloto.
 - Errores por el entrevistador:
 - Aumentar el número de entrevistadores.
 - Adecuar su selección a la encuesta y a la población a entrevistar.
 - Mejorar su formación.
 - Estandarizar su actuación.
 - Intensificar su supervisión y control.
 - Errores por el entrevistado:
 - Mejorar el diseño del cuestionario.
 - Mejorar su aplicación.
 - Errores por el tratamiento de información:
 - Automatización del procedimiento de encuesta.
 - Mejorar la formación y el control de los responsables.
 - Considerar el diseño muestral en el análisis estadístico, comprobar los supuestos estadísticos y adecuar el análisis a la naturaleza de los datos y tamaño muestral.
- Errores de subcobertura:
 - Evaluar la calidad del marco muestral.

- Restringir la población de estudio a la población marco.
 - Utilizar conjuntamente varios marcos muestrales.
 - Identificar las omisiones y utilizarlas como variables de cuota.
 - Aplicar otros métodos de encuesta.
 - Ajustes estadísticos postencuesta mediante la ponderación de la muestra.
- Errores de falta de respuesta:
- Adecuar el trabajo de campo a la población. Si fuese necesario, ampliar su duración.
 - Notificación antes de la encuesta.
 - Seguimiento de la muestra.
 - Gratificar la participación.
 - Mejorar la presentación de la encuesta.
 - Emplear otro método de encuesta en los no respondientes.
 - Convertir las “negativas” en aceptación.
 - Incrementar el tamaño de muestra.
 - Sustituir a los no respondientes por unidades equiparables.
 - Ajustes estadísticos postencuesta mediante la ponderación de la muestra y la imputación.

Las diferentes técnicas de calibración de estimadores presentadas en esta tesis doctoral pueden mejorar la representatividad de los datos, merceda por los errores ajenos al muestreo relacionados con subcobertura o falta de respuesta.

1.5 UTILIDAD DE LAS ENCUESTAS

Entre otras posibilidades, las encuestas aplicadas al ámbito de la Salud y Bienestar permitirían:

- conocer el estado y la evolución de las poblaciones de estudio en las diferentes etapas de la vida,
- utilizar la encuesta como marco muestral para la extracción de submuestras con el fin de ampliar información sobre grupos, poblaciones específicas o subpoblaciones con determinadas características (como enfermos crónicos, minorías étnicas, personas con necesidades especiales o con bajos ingresos, menores de edad, etc.),

- diseñar y evaluar políticas y programas de intervención en base a los resultados obtenidos,
- proporcionar información de causalidad en fenómenos complejos,
- incorporar temas de actualidad de una edición u oleada a otra en función de las necesidades de la población y de los sistemas de atención,
- proporcionar paneles de datos sobre determinantes de salud y de utilización de servicios sociosanitarios y
- complementar, contrastar o completar información de registros administrativos.

Las ventajas, posibilidades y utilidades de las encuestas comentadas en los apartados anteriores hacen necesario el uso de este tipo de herramientas, claves para el desarrollo y la mejora continua de los Sistemas de Salud.

Parte II

MÉTODOS DE ESTIMACIÓN INDIRECTOS

La *información auxiliar* es la proporcionada por las encuestas sobre variables relacionadas con la variable de estudio que, a su vez, también pueden ser obtenidas para la totalidad de la población de estudio. La forma de utilizar esta información auxiliar da lugar a diferentes métodos de estimación de los parámetros, haciéndolos más precisos y válidos. Esta segunda parte se divide en 3 capítulos. El [Capítulo 2](#) comienza con un desarrollo histórico de los primeros métodos utilizados para incorporar la información auxiliar en la fase de estimación, continúa explicando cuándo utilizarla y finaliza introduciendo los *estimadores indirectos* más utilizados. El [Capítulo 3](#) introduce el método de *verosimilitud empírica*, las modificaciones de su función (que da lugar por ejemplo al estimador de verosimilitud pseudo empírica modelo-calibrado) y su aplicación bajo diferentes diseño muestrales. Por último, el [Capítulo 4](#) introduce la *regresión no paramétrica*, sus ventajas y estimadores como el de modelo asistido o el de calibración no paramétrico. Este capítulo, así como esta segunda parte, finaliza con la utilización de esos estimadores no solo con información auxiliar cuantitativa, sino también cualitativa.

CUESTIONES GENERALES

2.1 INTRODUCCIÓN

En la teoría del muestreo, se ha estudiado en profundidad la estimación de parámetros lineales y cuadráticos, como son el total, la media o la varianza poblacionales. Se han desarrollado diseños muestrales cada vez más complejos con el fin de obtener estimadores más precisos de dichos parámetros (muestreos polietápicos, muestreos adaptativos, estratificación a posteriori,...). La mayoría de estos métodos únicamente observan el comportamiento de la variable objeto de estudio para realizar la estimación. A menudo, sucede que la variable o característica que queremos estudiar, denominada y , se encuentra fuertemente relacionada con una o varias variables o características auxiliares, denominadas $\mathbf{x} = (x_1, x_2, \dots, x_j)$ y cuyos valores son conocidos, son fáciles de obtener para todos los elementos de la población o el coste necesario para su conocimiento es relativamente bajo. En ese sentido, la *información auxiliar* se podría obtener a partir de:

- Registros o encuestas de gran tamaño muestral.
- Muestras grandes no probabilísticas.
- Muestras pequeñas probabilísticas.
- Muestras o censos relativos a la población de estudio en fechas anteriores.
- Poblaciones diferentes y altamente relacionadas con la de estudio.

En esas situaciones, cabe plantearse cómo utilizar esa información auxiliar \mathbf{x} para mejorar las estimaciones directas o expandidas (aquellas calculadas únicamente a partir de la muestra de la variable de estudio). Es decir, construir nuevos estimadores, denominados indirectos, que:

- para el mismo tamaño muestral tengan menor error de estimación, lo que implicaría mayor precisión en las estimaciones de los parámetros, o
- tengan el mismo error que los ya conocidos pero con un menor tamaño muestral, lo que produciría una disminución en el coste de la realización de la encuesta.

En el caso de no disponer de estos beneficios, rechazaremos el uso de estos estimadores indirectos debido a la complejidad de su diseño y puesta en práctica frente a los estimadores directos.

Destacar que el uso de información auxiliar no solamente se emplea en los estimadores indirectos, sino que también la empleamos, entre otros casos, para:

- La puesta en práctica del muestreo con probabilidades desiguales (en el método de total acumulado, Lahiri, Midzuno, etc.), donde gracias a una variable auxiliar x , podemos seleccionar las unidades que van a pertenecer a la muestra. Con el método de Lahiri, por ejemplo, usamos un esquema de muestreo en el que las probabilidades de selección de cada unidad son proporcionales a la característica adicional x .
- La formación de los estratos en un muestreo estratificado. El criterio de equipartición, por ejemplo, emplea la suma acumulada de las raíces cuadradas de las frecuencias de una variable auxiliar para la división de los distintos estratos.

2.2 DESARROLLO HISTÓRICO

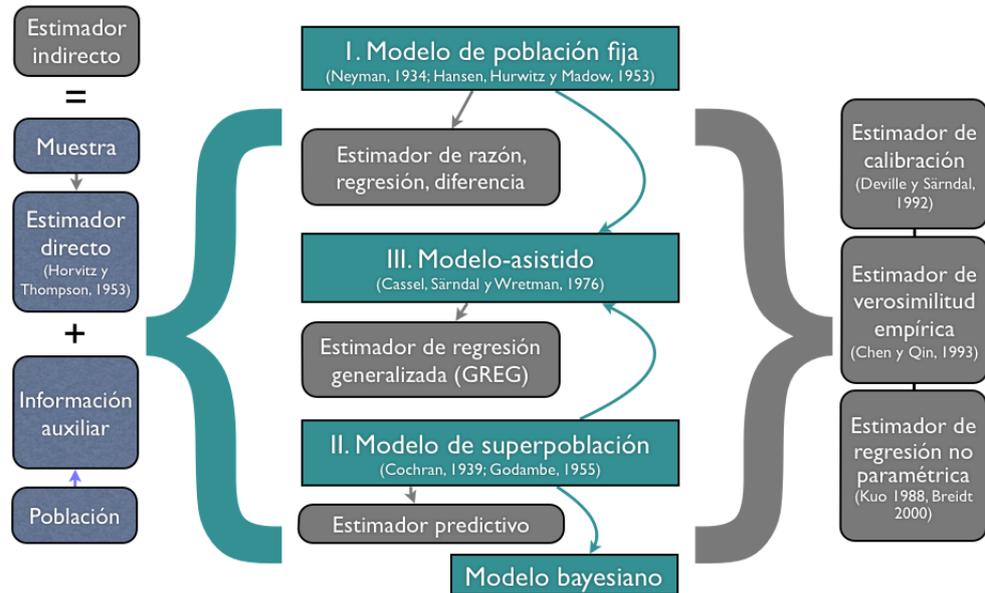
Así pues, los métodos de estimación indirectos consisten en modificar la forma del estimador directo con el fin de que el nuevo estimador gane en validez y precisión. Ese nuevo estimador, llamado indirecto, se construye con la información de la muestra y con otra información auxiliar que se dispone tanto en la muestra como en toda la población y que además está relacionada con la variable de estudio.

Estos estimadores indirectos no siempre garantizan que se produzca una disminución del error de muestreo respecto a los estimadores directos. La precisión que ganamos con los métodos que emplean la información auxiliar, estará en función del buen uso de las hipótesis que se usen para emplear un procedimiento u otro y en el que dichas hipótesis se ajusten, en mayor o menor medida, al problema real.

La propiedad más importante a destacar en los estimadores indirectos es que pueden abordarse desde los distintos enfoques hoy seguidos en el muestreo: modelo de población fija, modelo de superpoblación y modelo-asistido; siendo pues una metodología que puede ser aceptada por cualquiera de los estudiosos en muestreo, independientemente de su forma de hacer inferencia en este área. La [Figura 2.1](#) esquematiza esos enfoques y que, a continuación, se describen brevemente.

MODELO DE POBLACIÓN FIJA (O BASADO EN EL DISEÑO). Bajo este enfoque de la teoría clásica del muestreo en poblaciones finitas se

Figura 2.1: Métodos de estimación indirectos: enfoques



considera que los valores y_i de la característica de interés Y asociados a la i -ésima unidad de una población finita U son fijos aunque desconocidos, salvo para los elementos de la muestra una vez que ha sido obtenida. Por tanto, esos valores no tienen la consideración de aleatorios (no se consideran variables aleatorias). La aleatoriedad, bajo este enfoque, es fruto, exclusivamente, de la selección de la muestra y se refleja en el diseño muestral probabilístico adoptado y en los estimadores utilizados que introducen variables indicadoras de la pertenencia de una unidad a la muestra cuya distribución sólo depende del sistema de selección utilizado. Así pues, las inferencias están basadas en el diseño adoptado (R. Pérez, 2002). Los estimadores indirectos más conocidos bajo este enfoque son los de razón, regresión y diferencia.

Sin embargo, la inferencia bajo este enfoque no permite obtener estimadores óptimos o de mínima varianza por lo que ha recibido críticas de diversos autores, desarrollándose así otras aproximaciones a la inferencia en poblaciones finitas como la que se muestra a continuación.

MODELOS DE SUPERPOBLACIÓN (O ENFOQUE PREDICTIVO). Este enfoque fue introducido formalmente por Godambe (1955) para suplir la limitación del enfoque tradicional de no existencia de estimadores óptimos, aunque implícitamente se encuentra en trabajos anteriores sobre estimadores de regresión, muestreo sistemático y comparación de varianzas. Con la aparición de los modelos de superpoblación o enfoque predictivo (R. Pérez, 2002; G. Sánchez-Crespo, 2002; M. Thompson,

1988; Valliant y col., 2000) la teoría de muestreo tuvo un gran empuje pues se le dotó de un instrumento muy valioso que permitió obtener resultados más concluyentes en la comparación de estrategias y eventualmente producir estrategias óptimas en varias situaciones. En esta ocasión, los estimadores que se obtienen, llamados predictivos, son óptimos bajo ciertos modelos.

Supongamos por ejemplo que estudiamos el gasto en el consumo de las familias. Tenemos una población formada por N familias, numeradas de 1 a N . En lugar de centrarse en estudiar la variable en el momento de referencia, se puede considerar la naturaleza estocástica de la variable en estudio e intentar predecir el consumo en un determinado instante o a lo largo de un periodo.

Vemos los valores de la variable de interés como variables aleatorias Y_1, Y_2, \dots, Y_N que siguen una determinada función de distribución conjunta f . Haciendo esta consideración, los valores actuales y_1, y_2, \dots, y_N son una realización de las variables Y_1, Y_2, \dots, Y_N . Estos valores siguen siendo desconocidos a esos que se observen mediante una operación censal o de muestreo.

La distribución f es la que va a determinar el modelo de superpoblación. Éste resume el conocimiento que tengamos a priori de la naturaleza de la variable de estudio, basado en la experiencia o en argumentos subjetivos de nuestra percepción.

A través de un modelo de superpoblación se da una información más o menos completa que define en mayor o menor medida la clase de distribuciones a la que pertenece, puede ir desde dar los dos primeros momentos a establecer que sigue una distribución normal con medias y varianzas conocidas.

La aplicación de estos modelos requiere también del empleo de *variables auxiliares*, con sus valores conocidos para todos los individuos de la población. Concretamente los modelos de superpoblación más usados con gran diferencia son los modelos de regresión dados por:

$$\begin{aligned} y_k &= \mu(x_k) + \epsilon_k \quad \text{con} \quad E_{\xi}[\epsilon_k] = 0, \quad E_{\xi}[\epsilon_k^2] = \sigma^2 V(x_k) \quad \text{y} \\ E_{\xi}[\epsilon_j \epsilon_k] &= 0 \quad \forall k \neq j, \end{aligned} \quad (2.1)$$

donde ϵ_k es un error aleatorio.

Sin embargo, este enfoque también es criticado puesto que no tiene en cuenta cómo se ha seleccionado la muestra defendiendo que los estimadores basados en el diseño son más transparentes. Además, las estimaciones dependen del modelo elegido, dando estimaciones muy sesgadas cuando el modelo no es el correcto. Por lo que surge el siguiente otro enfoque.

MODELO-ASISTIDO (O INFERENCIA ASISTIDA POR EL MODELO). Este enfoque considera los anteriores dos enfoques, en el que [Cassel y col. \(1976\)](#); [Hedayat y Sinha \(1991\)](#) proponen los *estimadores de regresión generalizada*, que son óptimos si la población proviene de una superpoblación que sigue un modelo de regresión lineal. Esos estimadores pueden verse como una generalización del estimador de regresión donde el estimador del coeficiente de regresión usual se sustituye por un estimador ponderado por ciertos pesos que determina el modelo.

Así pues, esos enfoques nos permiten obtener otros estimadores en función de cómo se utilice la información auxiliar y así surgen los estimadores de calibración ([Parte III](#)), verosimilitud empírica ([Capítulo 3](#)) o de regresión no paramétrica ([Capítulo 4](#)), que describimos en los próximos capítulos. Algunos de esos métodos importantes para utilizar la información auxiliar, se han discutido solamente desde la perspectiva de los modelos de superpoblación, y concretamente usando un modelo de regresión. Además, aunque sus estimadores tienen muy buenas propiedades teóricas, pocas encuestas por muestreo hacen uso de estas metodologías.

A continuación describimos aquellos estimadores indirectos más utilizados bajo el enfoque de modelo de población fija.

2.3 ESTIMADORES INDIRECTOS CLÁSICOS

2.3.1 Definición del estimador general de regresión

Son métodos de estimación indirectos aquellos que usan una o varias variables auxiliares $\mathbf{x} = (x_1, x_2, \dots, x_J)$ altamente relacionadas con la característica objeto de estudio y , con el fin de obtener unos nuevos estimadores más precisos que los obtenidos únicamente con las observaciones de la muestra de la variable de estudio.

Los métodos de estimación indirecta más conocidos y utilizados son:

1. Estimación por razón, basado en la razón entre x e y .
2. Estimación por regresión, basado en la regresión entre x e y .
3. Estimación por diferencia, basado en la diferencia entre x e y .

No obstante, debemos tener en cuenta que no siempre obtendremos una mejor precisión con los estimadores indirectos y por tanto tendremos que hacer a priori un estudio de la hipótesis y de las propiedades de los datos que disponemos para emplear un método u otro. Así, por ejemplo,

si tenemos un tamaño de muestra grande, la precisión del estimador por razón es mejor que la obtenida a partir del estimador directo o de expansión simple, siempre que el coeficiente de correlación entre la variable adicional y la que es objeto de estudio sea positivo y alto. A su vez, la precisión del estimador de razón en esta situación se puede mejorar usando el estimador por regresión.

El marco de trabajo que usaremos a lo largo de este capítulo es el descrito brevemente a continuación:

La población U consta de N elementos distintos identificados a través de sus etiquetas $k = \{1, \dots, N\}$. Dentro de esta población estamos interesados en estudiar ciertas características de una variable de estudio que denominaremos y . Asociado al elemento k de la población podemos conocer exactamente y sin error el valor de la característica de interés, esta cantidad se denotará como y_k . Observar el valor y_k en todas las unidades de la población va a resultar imposible o muy costoso, así que utilizaremos una muestra para conocer los valores y_k de las unidades que pertenecen a la muestra. Una muestra es por tanto un subconjunto de n elementos ($n < N$), s , de U con sus valores asociados de y , es decir, $\{(k, y_k)\}$, seleccionados de acuerdo con un diseño de muestreo específico, $d(s)$, que asigna una probabilidad conocida $p(s)$ tal que $p(s) > 0$ para todo $s \in S$, conjunto de las posibles muestras s y $\sum_{s \in S} p(s) = 1$.

Suponemos que estamos interesados en estimar el total poblacional de la variable y , es decir, estimar

$$T_y = \sum_{k \in U} y_k, \quad (2.2)$$

y que x es una variable auxiliar que está correlacionada con la de interés, y .

Definición. El *estimador general de regresión* para el total poblacional de y se define como

$$\hat{T}_{yG} = \hat{T}_y + b_0(T_x - \hat{T}_x), \quad (2.3)$$

donde:

- \hat{T}_y y \hat{T}_x son los estimadores del total de y y de x , respectivamente, bajo el diseño muestral que deseemos. Por ejemplo, bajo muestreo aleatorio simple (MAS) los estimadores serían:

$$\hat{T}_{ymas} = \frac{N}{n} \sum_{i=1}^n y_i \quad y \quad \hat{T}_{xmas} = \frac{N}{n} \sum_{i=1}^n x_i.$$

Los estimadores de *Horvitz y D. Thompson (1953)* son:

$$\hat{T}_{yH} = \sum_{i=1}^n \frac{1}{\pi_i} y_i = \sum_{i=1}^n d_i y_i \quad y \quad \hat{T}_{xH} = \sum_{i=1}^n \frac{1}{\pi_i} x_i = \sum_{i=1}^n d_i x_i,$$

donde π_i son las probabilidades de primer orden tales que $\pi_i > 0$ para todo $i \in s$ y $d_i = \frac{1}{\pi_i}$.

- b_0 puede interpretarse como un coeficiente de corrección para mejorar el estimador \hat{T}_y .

Según el valor de b_0 tendremos los siguientes casos:

- *Estimador directo o expandido*. Lo obtendremos cuando $b_0 = 0$ y, por lo tanto, $\hat{T}_{yG} = \hat{T}_y$.
- *Estimador de razón o por cociente*. Es el que resulta cuando

$$b_0 = \frac{\hat{T}_y}{\hat{T}_x} = \hat{R},$$

obteniendo como estimador del total de y

$$\hat{T}_{yG} = \hat{T}_y + \frac{\hat{T}_y}{\hat{T}_x} (T_x - \hat{T}_x) = \hat{T}_{yR} = \frac{\hat{T}_y}{\hat{T}_x} T_x.$$

- *Estimador de regresión generalizada*. Con $b_0 = b =$ coeficiente de regresión de y_i sobre x_i obtenemos el estimador de regresión generalizada (**GREG**), \hat{T}_{yrg} ,

$$\hat{T}_{yG} = \hat{T}_{yrg} = \hat{T}_y + b(T_x - \hat{T}_x).$$

- *Estimador por diferencia*. Es el que resulta con $b_0 = 1$, siendo por tanto su expresión:

$$\hat{T}_{yG} = \hat{T}_{yD} = \hat{T}_y + T_x - \hat{T}_x.$$

En los siguientes apartados se desarrollan más esos estimadores.

2.3.2 Estimador de razón o por cociente

Cuando disponemos de una variable auxiliar x relacionada con la característica objeto de estudio y , nos plantearemos utilizar el método de

razón para tratar de obtener menor error de muestreo que el obtenido a través de un estimador de expansión simple.

El estimador de razón para el *total poblacional* de y es

$$\hat{T}_{yR} = \frac{\hat{T}_y}{\hat{T}_x} T_x = \frac{\bar{y}}{\bar{x}} T_x.$$

Si lo que estimamos es la *media* de y , el estimador de razón es

$$\hat{Y}_R = \frac{\bar{y}}{\bar{x}} \bar{X}.$$

En estos casos también puede resultar de interés estimar el *cociente* entre las dos variables, es decir,

$$R = \frac{T_y}{T_x}.$$

El estimador en dicho caso sería

$$\hat{R} = \frac{\hat{T}_y}{\hat{T}_x} = \frac{\bar{y}}{\bar{x}}.$$

Y, por tanto, podemos expresar los estimadores de razón para la media y el total en función de \hat{R} , es decir,

$$\hat{T}_y = \hat{R} T_x \quad ; \quad \hat{Y}_R = \hat{R} \bar{X}. \quad (2.4)$$

Obviamente, el estimador de razón en general no es insesgado. El sesgo viene dado por

$$B(\hat{R}) = E[\hat{R}] - R = -\frac{cov(\hat{R}, \bar{x})}{\bar{X}}. \quad (2.5)$$

Existe una cota superior para el estimador de razón:

$$\frac{|B(\hat{R})|}{\sigma_{\hat{R}}} \leq CV_{\hat{x}}.$$

Entonces, aunque estos estimadores no son insesgados, si no hay demasiada distorsión (coeficiente de variación, CV) de la muestra para la variable auxiliar, el sesgo del estimador por cociente es despreciable en comparación con su desviación típica. El mismo límite se aplica al sesgo de los estimadores de razón para el total y la media.

La expresión de sesgo antes comentada es poco práctica, de modo que solemos utilizar una expresión aproximada para $B(\hat{R})$:

$$B(\hat{R}) \cong \frac{1-f}{n\bar{X}^2}(RS_x^2 - \rho S_y S_x),$$

donde

- $f = \frac{n}{N}$ es la fracción de muestreo.
- S_y^2 y S_x^2 son las cuasivarianzas poblacionales de y y de x .
- ρ es el coeficiente de correlación lineal poblacional entre y y x .

Basándonos en la muestra, la estimación del sesgo del estimador de razón sería

$$\hat{B}(\hat{R}) \cong \frac{1-f}{n\bar{X}^2}(\hat{R}s_x^2 - \hat{\rho}s_y s_x).$$

Como estos estimadores son sesgados, la precisión se mide a través del *error cuadrático medio* (*ECM*). Debido a que no existe expresión exacta para el *ECM*, usamos la siguiente aproximación válida en muestras grandes:

$$ECM(\hat{R}) \cong \frac{1-f}{n\bar{X}^2} \frac{\sum_{i=1}^N (y_i - Rx_i)^2}{N-1}.$$

Usando las expresiones [Ecuación 2.4](#), es fácil obtener la fórmula para los estimadores de razón de la media y el total, puesto que $ECM(\hat{T}_{yR}) = ECM(\hat{R}T_x) = \hat{T}_x^2 ECM(\hat{R})$ y $ECM(\hat{Y}_R) = ECM(\hat{R}\bar{X}) = \bar{X}^2 ECM(\hat{R})$.

Una estimación del *ECM*, viene dada por

$$\widehat{ECM}(\hat{R}) \cong \frac{1-f}{n\bar{X}^2} \frac{\sum_{i=1}^n (y_i - Rx_i)^2}{n-1}.$$

Como ya hemos dicho, no siempre obtenemos mejor precisión con los estimadores de razón. Podemos analizar, por ejemplo, qué circunstancias han de ocurrir para que el estimador de razón tenga mejor precisión que el estimador bajo *MAS*. Para ello, debemos comparar los errores de muestreo de ambos estimadores, es decir, tenemos que ver cuándo $\sqrt{ECM(\hat{T}_{yR})} < \sqrt{ECM(\hat{T}_{ymas})}$, que despejando obtenemos que dicha desigualdad se cumple si y solamente si

$$\frac{1}{2} \leq \rho \frac{CV_x}{CV_y},$$

de donde se deduce que el estimador de razón es mejor que el de expansión simple en los siguientes casos:

- Si ρ es grande y positivo, esto es, si la variable auxiliar está muy correlacionada con la variable en estudio.
- Cuando el CV de la variable auxiliar sea pequeño, o lo que es lo mismo, los datos de la variable auxiliar sean muy homogéneos.

Todas las modificaciones del estimador de razón para el muestreo estratificado con probabilidades iguales o desiguales pueden verse en [Sukhatme y col. \(1984\)](#).

Por otro lado, si la regresión de y sobre x es lineal y pasa por el origen, esto es, $E(y|x) = \beta x$, entonces los estimadores de razón son insesgados. Esto es un caso particular, pero varios investigadores han buscado otros estimadores basados en la razón entre las variables y y x , que son insesgados. Entre ellos, el más destacado es el estimador de [Hartley y Ross \(1954\)](#). Su expresión es la siguiente:

$$\bar{r}_{HR} = \bar{r} + \frac{n(N-1)}{(n-1)N\bar{X}}(\bar{y} - \bar{r}\bar{x}),$$

donde

- $\bar{r} = \frac{1}{n} \sum_{i=1}^n \frac{y_i}{x_i}$
- \bar{r}_{HR} tiene la propiedad de insesgadez, es decir, $E(\bar{r}_{HR}) = R$

Los correspondientes estimadores del total y media poblacional son:

- $\hat{Y}_{HR} = \bar{r}_{HR}\bar{X}$
- $\hat{T}_{yHR} = \bar{r}_{HR}T_y$

Una aproximación de la varianza de este último estimador viene dada por la expresión

$$V(\hat{T}_{yHR}) \cong \frac{N^2}{n} \sum_{i=1}^N \frac{(y_i - \bar{Y} - \bar{r}_N(x_i - \bar{X}))^2}{N-1},$$

donde \bar{r}_N es la media poblacional de $r_i = \frac{y_i}{x_i}$.

La estimación de esta varianza viene dada por

$$\hat{V}(\hat{T}_{yHR}) \cong \frac{N^2}{n} \sum_{i=1}^n \frac{(y_i - \bar{y} - \bar{r}(x_i - \bar{X}))^2}{n-1}.$$

Otras transformaciones del estimador de razón para disminuir su sesgo o aumentar su precisión pueden verse en [J. Rao \(1994\)](#); [H. Singh y Espejo \(2003\)](#).

2.3.3 Estimador de regresión

Si la relación entre y y x es lineal y pasa por el origen, un estimador de razón es el apropiado, pero si la relación es lineal y no pasa por el origen, es recomendable utilizar la información adicional que proporciona x mediante un estimador de regresión.

El estimador de regresión viene dado por la expresión

$$\hat{T}_{yrg} = \hat{T}_y + b(T_x - \hat{T}_x),$$

donde b = coeficiente de regresión de y_i sobre x_i .

Se puede plantear este estimador de dos formas distintas, dependiendo de si b es conocido o no.

ESTIMADORES DE REGRESIÓN CON b CONOCIDO. Si b se conoce de antemano, bajo un **MAS** el estimador de regresión sería el anterior \hat{T}_{yrg} , que es insesgado con varianza

$$V(\hat{T}_{yrg}) = \frac{1-f}{n} (S_y^2 - 2bS_{yx} + b^2S_x^2) N^2.$$

Siendo su estimador

$$\begin{aligned} \hat{V}(\hat{T}_{yrg}) &= \frac{1-f}{n} (s_y^2 - 2bs_{yx} + b^2s_x^2) N^2 = \\ &= \frac{1-f}{n} \frac{1}{n-1} \sum_{i=1}^n [(y_i - \bar{y}) - b(x_i - \bar{x})]^2 N^2. \end{aligned}$$

ESTIMADORES DE REGRESIÓN CON b DESCONOCIDO. El estimador de regresión en este caso es el mismo que en el apartado anterior, pero b se calcula a partir de los datos de la muestra, esto es

$$b = \frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2}.$$

Para n suficientemente grande, la varianza del estimador es:

$$V(\hat{T}_{yrg}) = \frac{1-f}{n} S_y^2 (1 - \rho^2) N^2,$$

que puede estimarse por:

$$\begin{aligned}\widehat{V}(\widehat{T}_{yrg}) &= \frac{1-f}{n(n-2)} \sum_{i=1}^n [(y_i - \bar{y}) - b(x_i - \bar{x})]^2 N^2 = \\ &= \frac{1-f}{n(n-2)} \left[\sum_{i=1}^n (y_i - \bar{y})^2 + b^2 \sum_{i=1}^n (x_i - \bar{x})^2 \right] N^2.\end{aligned}$$

Analizando las varianzas del estimador de regresión y el de expansión simple, se llega a la conclusión de que el estimador de regresión siempre es más eficiente que el de expansión simple, excepto cuando $\rho^2 = 0$, en cuyo caso las estimaciones dan la misma precisión.

Al igual que los estimadores de razón, los estimadores de regresión se pueden aplicar bajo otros diseños muestrales. Estas modificaciones también están desarrolladas en [Sukhatme y col. \(1984\)](#).

2.3.4 Estimador de diferencia

Este estimador es más fácil de calcular que el estimador de razón o de regresión, aunque, en la mayoría de los casos, su precisión es peor respecto al resto de estimadores indirectos estudiados. Su expresión bajo [MAS](#) es la siguiente

$$\widehat{T}_{yD} = (\bar{y} + (\bar{X} - \bar{x}))N = (\bar{X} + \bar{d})N,$$

donde $\bar{d} = \bar{y} - \bar{x}$.

Este estimador también es insesgado, su varianza viene dada por:

$$V(\widehat{T}_{yD}) = \frac{1-f}{n} \sum_{i=1}^N \frac{(d_i - \bar{D})}{N-1} N^2,$$

siendo $d_i = y_i - x_i$ y $\bar{D} = \frac{1}{N} \sum_{i=1}^N d_i = \bar{Y} - \bar{X}$

Esta varianza se estima por

$$\widehat{V}(\widehat{T}_{yD}) = \frac{1-f}{n} \sum_{i=1}^n \frac{(d_i - \bar{d})}{n-1} N^2.$$

Todos estos métodos indirectos se pueden aplicar en cualquier diseño muestral (muestreo con probabilidades desiguales, muestreo estratificado, etc). Un desarrollo más amplio de estos estimadores pueden verse en [García-Fernández y Mayor-Gallego \(1995\)](#); [J. Rao \(1994\)](#); [Särndal, Swensson y J Wretman \(1992\)](#); [H. Singh y Espejo \(2003\)](#).

VEROSIMILITUD EMPÍRICA

3.1 INTRODUCCIÓN

En este capítulo ofrecemos una descripción del método de verosimilitud empírica (VE) en la estimación de la media y total de la población. Esta metodología fue usada por Owen (1988, 1990) como un método para la construcción de regiones de confianza con observaciones independientes. Owen afirmó que el ratio estadístico de VE tiene una distribución asintótica χ^2 , y por tanto se puede usar para la estimación de intervalos de confianza y contraste de hipótesis. Por otra parte, Qin y Lawless (1994, 1995) usan por primera vez el método de VE para la estimación puntual cuando la información se incorpora a través de la maximización de la función de VE. A partir de esa aplicación, este método se extendió y una gran gama de desarrollos sobre VE han sido descritos para distintos ámbitos (Owen, 2001).

Históricamente el uso de VE fue propuesto por Hartley y J. Rao (1968), pero su primera aplicación formalizada en muestreo para poblaciones finitas se debe a J. Chen y Qin (1993), que lo estudiaron inicialmente bajo MAS y posteriormente para un diseño muestral general (J. Chen y R Sitter, 1999).

3.2 EL MÉTODO DE VEROSIMILITUD EMPÍRICA

De forma introductoria, el método de VE consiste en lo siguiente:

Sea $\mathbf{x} = (x_1, x_2, \dots, x_j)$ un vector de variables auxiliares perfectamente conocido para todos los elementos de la población U . Estamos interesados, por ejemplo, en estimar la media poblacional de la característica de interés, y , es decir, estimar

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^N y_i.$$

Si consideramos MAS, el estimador para este parámetro es

$$\bar{y} = \frac{1}{n} \sum_{i \in s} y_i = \sum_{i \in s} \frac{1}{n} y_i.$$

Se observa que utilizamos los n puntos y_i de la muestra con el mismo peso ($1/n$) para estimar el parámetro. Puede ocurrir que ciertas

observaciones y_i sean más determinantes que otras para el cálculo del parámetro. Bajo estas circunstancias es conveniente darle a las observaciones más determinantes un mayor peso que aquellas que son menos influyentes para estimar el valor del parámetro. Esta es la idea de los estimadores de VE, es decir, pretenden cambiar los pesos $1/n$ por otros pesos \hat{p}_i , $i = \{1, \dots, n\}$, con el objetivo de mejorar la estimación del parámetro. Aquí es donde entran en juego las variables auxiliares x , ya que se usan para obtener estos nuevos pesos.

3.3 EL ESTIMADOR MÁXIMO VEROSIMIL EMPÍRICO

Sea p_i la masa de probabilidad de y_i , con $i \in s$. El estimador máximo verosimil empírico de \bar{Y} se define como

$$\hat{\bar{Y}} = \sum_{i \in s} \hat{p}_i y_i,$$

donde \hat{p}_i , $\{i = 1, \dots, n\}$, maximiza la función de VE, $L(\mathbf{p}) = \prod_{i \in s} p_i$ (J. Chen y Qin, 1993), sujeta a las restricciones

- $\sum_{i \in s} p_i = 1$ ($p_i > 0$) y
- $\sum_{i \in s} p_i x_i = \bar{X}$.

La información auxiliar se incorpora en la segunda restricción, basándose en la premisa de que si esos pesos dan una estimación perfecta para la media poblacional de la variable auxiliar x , deben ser también unos buenos pesos para estimar la media poblacional de y . Por esto, las estimaciones serán mejores cuanto mayor sea la relación entre x e y .

Este problema de maximización con restricciones se resuelve mediante el método de los *multiplicadores de Lagrange*, existiendo diversas formas de enfocar la resolución del procedimiento. Daremos su solución en las secciones siguientes cuando nos centremos en un diseño muestral específico.

3.4 MODIFICACIONES DE LA FUNCIÓN DE VEROSIMILITUD Y SUS RESTRICCIONES

El método de VE ofrece varias modificaciones debido a que:

1. La función de VE $L(\mathbf{p})$ se reemplaza por otra función.

Este cambio se debe porque es preferible trabajar con la *log-función de verosimilitud empírica* definida como

$$l(\mathbf{p}) = \log \prod_{i \in s} p_i = \sum_{i \in s} \log(p_i),$$

en lugar de la función de verosimilitud empírica $L(\mathbf{p})$. Así pues, maximizar $L(\mathbf{p}) = \prod_{i \in S} p_i$ es equivalente a maximizar $l(\mathbf{p}) = \sum_{i \in S} \log(p_i)$, expresión desde la que es más fácil calcular las estimaciones al utilizar sumatorias. El problema que se plantea ahora es estimar la verdadera función de VE $l(\mathbf{p}) = \sum_{i=1}^N \log(p_i)$ a través de una función $\hat{l}(\mathbf{p}) = \sum_{i=1}^n \log(p_i)$, que es un planteamiento similar a lo estudiado hasta ahora, si consideramos $\log(p_i)$ como una variable de la que se pretende estimar su total.

Para un diseño general, J. Chen y R Sitter (1999) proponen estimar $l(\mathbf{p}) = \sum_{i=1}^N \log(p_i)$ a través de la denominada *log-función de verosimilitud pseudo empírica*, esto es

$$\hat{l}(\mathbf{p}) = \sum_{i \in S} d_i \log(p_i),$$

donde d_i son pesos diseñados básicos que hacen que $\hat{l}(p)$ sea insesgada bajo el diseño para $l(p)$, es decir

$$E[\hat{l}(\mathbf{p})] = E\left[\sum_{i \in S} d_i \log(p_i)\right] = \sum_{i=1}^N \log(p_i) = l(\mathbf{p}).$$

Sin duda, este cambio en la función de VE es otra modificación del método original propuesto por J. Chen y Qin (1993) y que posee propiedades interesantes como una mejora de la robustez y de la eficiencia.

2. Se reemplaza la restricción $\sum_{i \in S} p_i \mathbf{x}_i = \bar{\mathbf{X}}$.

Al imponer que $\sum_{i \in S} p_i \mathbf{x}_i = \bar{\mathbf{X}}$, estamos considerando valores para p_i que proporcionan estimaciones perfectas para $\bar{\mathbf{X}}$, y nos podemos plantear cómo de efectivo es el uso que estamos haciendo de la información adicional a través de la condición anterior. Por este motivo, si la información auxiliar $\mathbf{x} = (x_1, x_2, \dots, x_J)$ es conocida, una cuestión a preguntarse sería: *¿Cuál es la mejor expresión a usar en las restricciones para hallar el estimador de verosimilitud empírica?* Para resolver esta pregunta definiremos $\mathbf{u}_i = u(y_i, \mathbf{x}_i)$, con $i = \{1, \dots, N\}$, siendo $u(\cdot)$ una función (de calibración) conocida de y_i y de \mathbf{x}_i y que verifica

$$\frac{1}{N} \sum_{i=1}^N \mathbf{u}_i = \mathbf{0}. \quad (3.1)$$

Usaremos \mathbf{u}_i como una variable de calibración y reemplazaremos la expresión, $\sum_{i \in s} p_i \mathbf{x}_i = \bar{\mathbf{X}}$ por

$$\sum_{i \in s} p_i \mathbf{u}_i = \frac{1}{N} \sum_{i=1}^N \mathbf{u}_i = \mathbf{o}.$$

El planteamiento que surge ahora es cómo escoger \mathbf{u}_i para obtener estimadores más eficientes.

Este método origina infinidad de modificaciones o de soluciones dependiendo de la función \mathbf{u}_i escogida.

3. Se supone que la relación entre la variable de interés y y el vector de variables auxiliares $\mathbf{x} = (x_1, \dots, x_j)$ puede ser descrita a través de un *modelo de superpoblación*.

En este sentido, se han propuesto dos modelos que dan lugar a los estimadores óptimos modelos-asistidos y a la aproximación modelo-calibrada. Los primeros usan el criterio de mínima esperanza bajo el modelo de superpoblación de la varianza basada en el diseño. Ver por ejemplo los trabajos de [Cassel y col. \(1976\)](#); [Godambe \(1955\)](#); [Godambe y M. Thompson \(1973\)](#). Suponemos que y_1, y_2, \dots, y_N es una MAS de un modelo de superpoblación tal que

$$E_{\xi}(y_i) = \mu_i, \quad V_{\xi}(y_i) = \sigma_i^2, \quad i = 1, 2, \dots, N,$$

y y_1, y_2, \dots, y_N son independientes entre ellos. E_{ξ} y V_{ξ} denota la esperanza y la varianza bajo el modelo de superpoblación. En [C Wu \(2003\)](#) se demuestra que usando $\mathbf{u}_i = \mu_i$ como variable de calibración obtenemos un estimador con menor $E_{\xi}[AV_p(\hat{Y}_{PE})]$ entre todos los posibles valores $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N)$ tal que

$$\frac{1}{N} \sum_{i=1}^N (\mathbf{u}_i - \bar{\mathbf{u}})^2 \rightarrow \mathbf{c} \neq \mathbf{o}.$$

AV_p denota la varianza asintótica bajo el diseño.

3.5 APLICACIÓN A DIFERENTES DISEÑOS MUESTRALES

La metodología de [VE](#) es usada para la construcción de cualquier parámetro lineal (\bar{Y}, T_y, \dots) , función de distribución, cuantiles,..., y se puede aplicar bajo cualquier diseño muestral ([MAS](#), muestreo con probabilidades desiguales, estratificado, por conglomerados,...). Para ello, a priori debemos:

1. Determinar la función $u(\cdot)$ utilizada en las restricciones.

2. Enfocar el problema bajo un modelo de población fija, es decir, basado en el diseño muestral o basándonos en la hipótesis de un modelo de superpoblación, usando el enfoque modelo-asistido o modelo-calibrado.

Una vez tenidas en cuenta estas consideraciones previas, la aplicación del método de VE se extiende a cualquier diseño muestral sin más que asignarle a los pesos diseñados básicos comentados anteriormente, el valor que le corresponda en función del esquema de muestreo seleccionado.

A continuación se muestran aplicaciones a algunos diseños muestrales:

MUESTREO ALEATORIO SIMPLE. Bajo un diseño MAS, con probabilidades iguales y sin reemplazamiento, tenemos que

$$\pi_i = \frac{n}{N}$$

$$d_i = \frac{1}{\pi_i} = \frac{N}{n}$$

$$\sum_{j=1}^n d_j = N,$$

donde $i = \{1, \dots, n\}$ y π_i son las probabilidades de inclusión de primer orden de la característica y_i , es decir, la probabilidad que tiene la i -ésima unidad de la población de pertenecer a la muestra bajo un diseño muestral dado, en este particular, MAS.

Y así,

$$d_i^* = \frac{d_i}{\sum_{j=1}^n d_j} = \frac{1}{n}.$$

La expresión que tenemos para el estimador de \bar{Y} es $\hat{Y} = \sum_{i=1}^n \hat{p}_i y_i$ con

$$\hat{p}_i = \frac{1}{n(1 + \lambda' u_i)}.$$

Con u_i definido como en [Ecuación 3.1](#) y λ como solución, por el método de multiplicadores de Lagrange, de la ecuación

$$\sum_{i=1}^n \frac{1}{1 + \lambda' u_i} = 0.$$

De manera que este estimador coincide con el propuesto por [J. Chen y Qin \(1993\)](#).

MUESTREO CON PROBABILIDADES IGUALES Y REEMPLAZAMIENTO. En los diseños muestrales con reemplazamiento se tiene que,

$$d_i = \frac{1}{n\alpha_i}.$$

Siendo α_i la probabilidad de que la i -ésima unidad sea seleccionada para la muestra.

Además, en un muestreo con probabilidades iguales se cumple que,

$$\alpha_i = \frac{1}{N'}$$

y, por tanto,

$$d_i = \frac{N}{n}.$$

Coincidiendo con los pesos de un [MAS](#), con lo que se usan las mismas expresiones que en ese diseño, pero con la diferencia de que es posible que la muestra contenga unidades repetidas.

MUESTREO CON PROBABILIDADES DESIGUALES Y SIN REEMPLAZAMIENTO. Para este diseño muestral usaremos

$$d_i = \frac{1}{\pi_i},$$

donde $i = \{1, \dots, n\}$ y π_i son las probabilidades de inclusión de primer orden de la característica y_i .

El estimador de verosimilitud empírica que resulta para \bar{Y} será

$$\widehat{\bar{Y}} = \sum_{i \in s} \hat{p}_i y_i,$$

con

$$\hat{p}_i = \frac{\frac{d_i}{\sum_{j=1}^n d_j}}{1 + \lambda' u_i} = \frac{d_i^*}{1 + \lambda' u_i},$$

donde u_i se define como en [Ecuación 3.1](#) y λ como solución por el método de los multiplicadores de Lagrange de la ecuación,

$$\sum_{i=1}^n \frac{\frac{d_i}{\sum_{j=1}^n d_j} u_i}{1 + \lambda' u_i} = 0.$$

MUESTREO CON PROBABILIDADES DESIGUALES Y CON REEMPLAZAMIENTO. En los diseños muestrales con reemplazamiento,

$$d_i = \frac{1}{n\alpha_i},$$

donde α_i es la probabilidad de que la i -ésima unidad sea seleccionada para la muestra. Y, entonces,

$$d_i^* = \frac{1/\alpha_i}{\sum_{j=1}^n 1/\alpha_j}.$$

Resolviendo el estimador como anteriormente,

$$\hat{Y} = \sum_{i \in s} \hat{p}_i y_i,$$

con

$$\hat{p}_i = \frac{\frac{1}{\alpha_i \sum_{j=1}^n 1/\alpha_j}}{1 + \lambda' u_i},$$

donde u_i se define como en [Ecuación 3.1](#) y λ como solución por el método de los multiplicadores de Lagrange de la ecuación,

$$\sum_{i=1}^n \frac{\frac{u_i}{\alpha_i \sum_{j=1}^n 1/\alpha_j}}{1 + \lambda' u_i} = 0.$$

MUESTREO ESTRATIFICADO. Vamos a hacer extensiva la metodología empleada a un tipo de muestreo más complejo, como es el muestreo estratificado. Un MAS proporciona mejores o peores resultados en función del tamaño de la muestra y la heterogeneidad de la población de estudio. Dado que, aumentar el tamaño muestral aumenta también los costes, una alternativa para mejorar la precisión es reducir la heterogeneidad. En esto se basa el muestreo estratificado. Sea una población de tamaño N , dividida en subpoblaciones de tamaño N_1, N_2, \dots, N_L , tal que $\sum_{h=1}^L N_h = N$. Estas subpoblaciones reciben el nombre de *estratos*.

Se selecciona una muestra en cada estrato, siendo la muestra final el conjunto de todas ellas. La selección dentro de cada estrato es independiente del resto, pudiendo elegir el tipo de muestreo más apropiado en cada estrato. Este tipo de muestreo es muy común, y especialmente recomendable cuando la población es heterogénea. Esta se divide en subpoblaciones lo más homogéneas posibles, mejorando así la precisión.

La notación es la siguiente:

N_h : número de unidades o tamaño poblacional del estrato h .

n_h : tamaño muestral en el estrato h .

$W_h = N_h/N$: ponderación del estrato h .

$f_h = n_h/N_h$: fracción de muestreo en el estrato h .

\bar{X}_h : media poblacional en el estrato h .

\bar{x}_h : media muestral en el estrato h .

S_h^2 : cuasivarianza poblacional en el estrato h .

s_h^2 : cuasivarianza muestral en el estrato h .

Así pues, el estimador directo lineal e insesgado para la media poblacional se expresa como

$$\widehat{Y}_{st} = \sum_{h=1}^L W_h \bar{y}_h.$$

En este contexto, se define la *log-función de verosimilitud empírica en muestreo estratificado* como

$$l(\mathbf{p}) = \sum_{h=1}^L \sum_{i=1}^{N_h} \log(p_{ih}).$$

Si la observamos como un total poblacional, la estimación insesgada a partir de una muestra y bajo un diseño muestral específico, viene dada por la denominada *log-función de verosimilitud pseudo empírica*,

$$\hat{l}(\mathbf{p}) = \sum_{h=1}^L \sum_{i=1}^{n_h} d_{ih} \log(p_{ih}),$$

de forma que los pesos d_{ih} hacen que $\hat{l}(p)$ sea insesgada bajo el diseño dado para $l(p)$.

Si tomáramos como diseño muestral dentro de cada estrato un MAS, lo que se denomina como muestreo estratificado aleatorio, los pesos resultan

$$d_{hi} = \frac{N_h}{n_h}.$$

El estimador de verosimilitud pseudo empírica se obtiene maximizando $\hat{l}(p)$ sujeta a las restricciones:

$$\sum_{i=1}^{n_h} p_{ih} = 1 \quad (p_i > 0) \quad \text{y} \quad \sum_{h=1}^L W_h \sum_{i=1}^{n_h} p_{ih} \mathbf{x}_{ih} = \bar{\mathbf{X}}.$$

Por comodidad, en la segunda restricción se ha considerado una relación lineal entre Y y X . Sin embargo, es posible modificarla en el caso de querer tomar otro tipo de relación entre las variables.

Conocidas ya las soluciones \hat{p}_{ih} del problema, el *estimador de verosimilitud pseudo empírica bajo muestreo estratificado* viene dado por,

$$\hat{Y}_{VPEst} = \sum_{h=1}^L W_h \sum_{i=1}^{n_h} \hat{p}_{ih} \mathbf{Y}_{ih}.$$

Ahora bien, dependiendo de si las medias poblacionales de la variable auxiliar dentro de cada estrato son conocidas o no, es decir, \bar{X}_h conocidas o no, el cálculo del estimador se puede plantear de formas diferentes:

- \bar{X}_h conocidas, y suponiendo una relación lineal, la segunda restricción se puede sustituir por la expresión

$$\sum_{h=1}^L W_h \sum_{i=1}^{n_h} p_{ih} \mathbf{x}_{ih} = \bar{\mathbf{X}} \approx \sum_{i=1}^{n_h} p_{ih} \mathbf{x}_{ih} = \bar{\mathbf{X}}_h,$$

y el cálculo del estimador de verosimilitud pseudo empírica es muy sencillo, solo hay que calcularlo dentro de cada estrato, \hat{Y}_{VPEst} , y el estimador final tiene la expresión

$$\widehat{Y}_{VPEst} = \sum_{h=1}^L W_h \widehat{Y}_{VPEh}.$$

- \bar{X}_h desconocidas, entonces se complica encontrar una solución. J. Chen y R Sitter (1999) propusieron una estrategia que consistía en considerar una restricción arbitraria para cada estrato en lugar de la segunda restricción, buscar la solución intermedia bajo esta situación y obteniendo la solución final mediante verosimilitud empírica. Otra estrategia debida a C Wu (2004) consiste en reemplazar las restricciones de manera que las nuevas estén todas formuladas al nivel del conjunto de los estratos. Ambas propuestas pueden ser consultadas en los textos originales o en F. Muñoz (2006).

OTROS DISEÑOS. La aplicación del método de verosimilitud empírica para el muestreo por conglomerados y muestreo bietápico estratificado, se puede consultar en J. Chen y R Sitter (1999, págs. 400 y 393-394, respectivamente).

REGRESIÓN NO PARAMÉTRICA

4.1 INTRODUCCIÓN

Los métodos de regresión no paramétricos y semiparamétricos son herramientas estadísticas que han ganado aceptación en la mayoría de los ámbitos de la Estadística. Hacen posible analizar los datos, hacer predicciones y realizar inferencias sin tener que especificar completamente un modelo paramétrico para los datos. Las técnicas de estimación no paramétricas empiezan a emplearse a partir del año 1940, pero no es hasta el año 1980 cuando su expansión y uso empiezan a hacerse patentes. Publicaciones como las de [Fan y Gijbels \(1996\)](#); [Härdle y Marron \(1991\)](#); [Wand y Jones \(1995\)](#), entre otras, son de gran importancia en el desarrollo de esta nueva metodología.

Se pueden citar un gran número de ventajas de los estimadores no paramétricos, así [Härdle y Marron \(1991\)](#) menciona como puntos a destacar los siguientes:

- proporcionan un método flexible para explorar una relación general entre las variables,
- permiten predecir observaciones sin considerar un modelo paramétrico fijo,
- proporcionan una herramienta para encontrar datos atípicos y
- constituyen un método flexible de sustitución para valores perdidos e interpolación de valores adyacentes de la variable explicativa.

En este capítulo, nos centraremos en su aplicación a la estimación de medias y totales en poblaciones finitas. En este sentido, los métodos no paramétricos se han introducido recientemente en marcos muestrales con poblaciones finitas. En el caso especial de los métodos basados en Kernel, [Kuo \(1988\)](#) propone un estimador de la función de distribución basado en el modelo, utilizando la regresión polinómica. [Breidt y Opsomer \(2000\)](#) y [Opsomer y Miller, \(2005\)](#) introdujeron un estimador modelo-asistido y un selector de ancho de banda (o parámetro de suavizamiento), respectivamente, mediante la incorporación del diseño muestral al Kernel.

Como se introdujo en el [Sección 2.2](#), mediante el uso de la información auxiliar, una variedad de enfoques están disponibles para la construcción de estimadores más eficientes para las medias y totales poblacionales, incluidos los métodos basados en el diseño y basados en el modelo (véase, por ejemplo, [Särndal, Swensson y JH Wretman 1989](#); [Valliant y col. 2000](#)). Según [RJA Little \(2004\)](#), el muestreo quizás sea el único en el que el área más importante de su actividad estadística actual es aquella donde las inferencias se basan principalmente en la distribución en el muestreo y no en modelos estadísticos de la variable de interés.

4.2 ENFOQUES BASADOS EN MODELOS Y MODELO-ASISTIDO

El enfoque de modelos se basa en los de superpoblación, que asumen que la población objeto de estudio $\mathbf{y} = (y_1, \dots, y_N)'$ es un conjunto de variables aleatorias en un modelo de superpoblación ξ . Para incorporar la información auxiliar \mathbf{x}_j disponible para todos los $j \in U$, se puede asumir un modelo de superpoblación para y construido sobre alguna función media de \mathbf{x} :

$$y_j = m(\mathbf{x}_j) + e_j, \quad j = 1, \dots, N. \quad (4.1)$$

El vector aleatorio $e = (e_1, \dots, e_N)'$ se asume con media cero y matriz de covarianzas definida positiva que es diagonal (y_j son mutuamente independientes).

Normalmente se utiliza el método paramétrico para representar la relación $m(\cdot)$ entre la información auxiliar y la variable de estudio. Sin embargo, la asignación de tal relación puede resultar inapropiada o no verificable. Una alternativa natural sugerida inicialmente por [Kuo \(1988\)](#) para la función de distribución, es adoptar una aproximación basada en un modelo no paramétrico, que no impone restricciones en la relación entre la información auxiliar y la variable de estudio.

Una vez el modelo es ajustado a los datos de la muestra, hay al menos dos maneras de incorporar estas predicciones dentro de la estimación del total de la población finita. La primera es mediante un enfoque basado en el modelo, en el que los valores del modelo ajustado se emplean para predecir sólo los valores no muestreados de y . La segunda forma es mediante el enfoque modelo-asistido, que evita problemas potenciales de mala especificación del modelo. La estimación modelo-asistida se basa en la predicción ([Ecuación 4.1](#)) para todos los elementos de la población pero corrigiendo el sesgo del diseño en esta predicción.

4.2.1 Estimador modelo-asistido

Breidt y Opsomer (2000) utiliza el estimador tradicional de regresión polinómica para la función desconocida $m(\cdot)$. Sea $K_h(u) = h^{-1}K(u/h)$, donde K es una función Kernel continua y h es el ancho de banda. Entonces, una predicción consistente para $m(x_j)$ es dada por

$$\hat{m}_j = e_1'(X_{sj}'W_{sj}X_{sj})^{-1}X_{sj}'W_{sj}Y_s = w_{sj}Y_s, \quad (4.2)$$

donde $e_1 = (1, 0, \dots, 0)'$ es un vector de longitud $p + 1$, donde p es el grado del polinomio, $Y_s = \{y_i\}_{i \in s}$, $W_{sj} = \text{diag}\{K_h(x_i - x_j)\}_{i \in s}$ y $X_{sj} = \{1, (x_i - x_j), \dots, (x_i - x_j)^p\}_{i \in s}$. Así los autores definen un estimador modelo-asistido de la media poblacional como

$$\hat{Y}_{BO} = \frac{1}{N} \left(\sum_{j \in s} \pi_j^{-1} (y_j - \hat{m}_j) + \sum_{j \in U} \hat{m}_j \right). \quad (4.3)$$

4.2.2 Estimador basado en el modelo

Este estimador es bien conocido por sus buenas propiedades teóricas y prácticas. Un estimador basado en el modelo es definido por Rueda e I Sánchez-Borrego (2009) como

$$\hat{Y}_{MB} = \frac{n}{N} \sum_{j \in s} y_j + \left(1 - \frac{n}{N}\right) \frac{1}{N - n} \sum_{j \in U - s} \hat{m}_j. \quad (4.4)$$

\hat{Y}_{MB} requiere que conozcamos los valores de x para todas las unidades de la población y cálculos intensivos y tiene las siguientes propiedades:

- Es lineal en el Y_s .
- No utiliza las probabilidades de diseño π_i . Como los nuevos pesos las ignoran la inferencia se hace de acuerdo con el principio de condicionalidad: la inferencia debería ser realizada condicionalmente a la muestra observada s , no promediada sobre todas las muestras que podrían haber sido seleccionadas.
- Es asintóticamente insesgada.

Los autores comparan, en la teoría y en la práctica, este estimador con el estimador modelo-asistido propuesto por Breidt y Opsomer (2000). Concluyen que ambas formas de usar modelos no paramétricos parecen plausibles, y que se pueden adoptar ambas filosofías del análisis estadístico, basada en el modelo y modelo-asistida, de acuerdo con las condiciones prácticas.

4.3 ESTIMADOR DE CALIBRACIÓN NO PARAMÉTRICO

Montanari y Ranalli (2005) combinan la estimación de calibración con métodos no paramétricos y proponen estimadores de calibración no paramétricos para la media de una población finita. Los autores adoptan una aproximación modelo-asistida para inferir y extender la calibración, como la de Changbao Wu y Randy R Sitter (2001), utilizando métodos no paramétricos para obtener los valores ajustados en los que calibrar. Así pues, asumen acceso a nivel unitario de la información auxiliar y consideran el aprendizaje de redes neuronales y suavizado polinomial para estimar la relación funcional entre la variable de interés y las auxiliares. Realizan diversos experimentos de simulación para comprobar el comportamiento en muestras finitas de los estimadores propuestos, y consideran un estimador modelo calibrado no paramétrico para evaluar la condición ecológica de corrientes en las tierras altas del Atlántico medio.

4.4 ESTIMADOR NO PARAMÉTRICO CON DISCONTINUIDADES

Rueda, I Sánchez-Borrego y A (2008) proponen un nuevo estimador no paramétrico basado en una versión adaptada del estimador Kernel lineal propuesto por IR Sánchez-Borrego y col. (2006) para estimar la función de regresión discontinua. Este método es el resultado de combinar el procedimiento de observaciones proyectadas, introducido por J. Wu y Chu (1993), conjuntamente con la regresión Kernel lineal, introducida por Fan y Gijbels (1996). El método propuesto se basa en dos pasos, inicialmente se estiman los puntos de salto y, posteriormente, la función de regresión junto con la media poblacional son estimados utilizando esos puntos de salto estimados.

Las observaciones proyectadas es un método de reutilización de los datos disponibles de cara a añadir nuevas observaciones en la región de localización del punto de salto, contribuyendo a mejorar la estimación de los puntos de salto. Denotamos esos estimadores como \hat{t}_k ($k = 1, \dots, q$), y consideramos que $\hat{t}_{k-1} \leq \hat{t}_k$, para $k = 1, \dots, q + 1$, definiendo $t_0 = 0$ y $t_{q+1} = 1$. La idea básica para estimar la función de regresión discontinua consiste en utilizar el estimador Kernel lineal en intervalos donde tal función es continua, es decir, intervalos del tipo $[\hat{t}_{k-1}, \hat{t}_k]$, y utilizando los datos proyectados para mejorar la estimación Kernel lineal a partir de nuevas observaciones en la región límite de tales intervalos.

El procedimiento se define para un diseño muestral general y es válido para estimar la media población finita cuando la función de regresión subyacente tiene más de una discontinuidad. Por otra parte, el

método es simple para ser implementado y se basa en un suavizado Kernel lineal cuyas buenas propiedades son bien conocidas en el contexto general de regresión no paramétrica.

4.5 VARIABLES AUXILIARES CUANTITATIVAS Y CUALITATIVAS

Todos esos estimadores no paramétricos consideran únicamente variables auxiliares cuantitativas. I Sánchez-Borrego y col. (2014) plantean una situación más general que incluye tanto información auxiliar cuantitativa como cualitativa.

Los autores asumen $\mathbf{x}_j = (x_{1j}, \dots, x_{kj})$ como los valores correspondientes a un vector \mathbf{x} conteniendo q variables auxiliares discretas y p continuas, con $q + p = k$. Sea \mathbf{x}^d un subvector de variables discretas q y \mathbf{x}^c el restante subvector de las continuas. Utilizamos x_t^c para el t componente de \mathbf{x}^c y x_t^d para el t componente de \mathbf{x}^d , respectivamente, y asumimos que x_t^d toma $c_t \geq 2$ valores diferentes en $D_t = \{0, 1, \dots, c_t - 1\}$.

Para los regresores discretos x_t^d , con $t = 1, \dots, q$, utilizamos una variación de la función Kernel de Aitchison y Aitken (1976), definida como

$$l_\lambda(x_{ti}^d - x_{tj}^d) = \begin{cases} 1 & \text{si } x_{ti}^d = x_{tj}^d \\ \lambda & \text{si } x_{ti}^d \neq x_{tj}^d \end{cases}, \quad (4.5)$$

donde λ es el parámetro de suavizado que satisface que $0 \leq \lambda \leq 1$, con $\lambda = 0$ correspondiendo a una función indicador y $\lambda = 1$ dando los mismos pesos a todos los valores de su argumento. El producto Kernel para las variables discretas se define como

$$K_{ij}^d = \prod_{t=1}^q l_\lambda(x_{ti}^d - x_{tj}^d). \quad (4.6)$$

Para las variables continuas, utilizamos K para denotar una función de densidad simétrica, univariante y define el producto Kernel como

$$K_{ij}^c = \prod_{t=1}^p \frac{1}{h^p} K\left(\frac{x_{ti}^c - x_{tj}^c}{h}\right), \quad (4.7)$$

donde h es el parámetro de suavizado y $0 < h < \infty$.

El producto Kernel para el caso de datos mixtos es simplemente el producto de K_{ij}^d y K_{ij}^c , esto es, el resultado de combinar (4.6) y (4.7),

$$\mathcal{K}_{ij} = \left[\prod_{t=1}^q l_\lambda(x_{ti}^d - x_{tj}^d) \right] \left[\prod_{t=1}^p \frac{1}{h^p} K\left(\frac{x_{ti}^c - x_{tj}^c}{h}\right) \right]. \quad (4.8)$$

Se propone el siguiente estimador modelo-asistido no paramétrico para $\hat{Y}_N = \sum_{j \in U} y_j / N$, la media poblacional:

$$\hat{Y}_{np} = \frac{1}{N} \sum_{j \in U} \hat{m}_j + \frac{1}{N} \sum_{j \in s} \frac{y_j - \hat{m}_j}{\pi_j}, \quad (4.9)$$

donde

$$\hat{m}_j = \frac{\sum_{i \in s} \mathcal{K}_{ij} y_i / \pi_i}{\sum_{i \in s} \mathcal{K}_{ij} / \pi_i}. \quad (4.10)$$

Este estimador es similar al de [Breidt y Opsomer \(2000\)](#), pero el estimador no paramétrico ahora permite tanto variables auxiliares continuas como categóricas.

El desarrollo para el estimador del total poblacional se haría multiplicando por N el de la media poblacional y su varianza multiplicando por N^2 .

Parte III

CALIBRACIÓN DE ESTIMADORES

Se trata de una de las más recientes técnicas para hacer más válida y precisa la estimación de un parámetro. Ofrece un marco general sobre ponderación con una *aproximación simple y práctica* para incorporar información auxiliar a la estimación. Esta tercera parte desarrolla el estado del arte de las técnicas de calibración y se compone de 3 capítulos. El [Capítulo 5](#) introduce cómo construir un *estimador de calibración*, sus condiciones y diferentes estimadores en función de la distancia elegida así como del proceso de construcción. El [Capítulo 6](#) desarrolla el caso de la calibración aplicada a *marcos imperfectos*, es decir, cuando se dan errores de cobertura. Así pues, se describen los diferentes tipos de esos errores y cómo reducirlos para los casos de uno o múltiples marcos imperfectos a través de los estimadores de calibración. Por último, el [Capítulo 7](#) introduce algunos de los programas informáticos específicos para la técnica de calibración de estimadores así como los paquetes relacionados con calibración del Software libre R.

CUESTIONES GENERALES

5.1 INTRODUCCIÓN

El método de calibración fue descrito por *Deville y Särndal (1992)*. Su objetivo principal es usar la información auxiliar para obtener estimadores aproximadamente insesgados y con varianza más pequeña que el estimador de *Horvitz y D. Thompson (1953)*. Para ello, la calibración requiere el conocimiento de los totales poblacionales de las variables auxiliares que se va a utilizar en el proceso. La eficiencia de un estimador de calibración depende de cómo de bien las variables auxiliares expliquen la variable de interés u objetivo.

En concreto, el método de calibración, consiste en sustituir los pesos (o factores de elevación) del estimador de *Horvitz-Thompson*, por unos nuevos pesos, de forma que estos pesos cumplan dos condiciones esenciales:

- Que los estimadores con estos nuevos pesos, den estimaciones perfectas, si se aplican a las variables auxiliares implicadas en el proceso de construcción del estimador de calibración.
- Que estén tan próximos como sea posible a los pesos del estimador de *Horvitz-Thompson*, de acuerdo con una métrica o distancia previamente definida.

Así pues, el método de calibración parte de dos ideas intuitivas:

1. Si se estima correctamente la variable auxiliar, lo mismo ocurrirá con la variable de interés.
2. Si no se estima bien la sencilla variable auxiliar, a saber qué pasará con la variable de interés.

5.2 CONSTRUCCIÓN DE UN ESTIMADOR DE CALIBRACIÓN

De forma muy breve, el marco de trabajo usual en muestreo de poblaciones finitas es el descrito a continuación.

Una población U consiste en N elementos distintos identificados a través de sus etiquetas $k = \{1, \dots, N\}$. La característica de interés y_k asociada con el elemento k se conoce exactamente (sin error) observando el elemento k . Una muestra es un subconjunto, s , de U y sus valores

asociados de y , es decir, $\{(k, y_k)\}$, seleccionados de acuerdo con un diseño de muestreo específico, $d(s)$, que asigna una probabilidad conocida $p(s)$ tal que $p(s) > 0$ para todo $s \in S$, conjunto de las posibles muestras s , y $\sum_{s \in S} p(s) = 1$.

Supongamos que queremos estimar el total poblacional de la variable y , es decir, queremos estimar

$$T_y = \sum_{k \in U} y_k.$$

Entonces, si consideramos el estimador de *Horvitz y D. Thompson (1953)* para T_y , es decir,

$$\hat{T}_{yH} = \sum_{k \in s} \frac{1}{\pi_k} y_k = \sum_{k \in s} d_k y_k,$$

donde π_k son las probabilidades de primer orden tales que $\pi_k > 0$ para todo $k \in s$ y $d_i = \frac{1}{\pi_i}$, así como un vector de J variables auxiliares, $\mathbf{x} = (x_1, \dots, x_J)$, de forma que es perfectamente conocido para todos los elementos de la población U , lo que pretendemos es modificar los pesos $d_k = \frac{1}{\pi_k}$, por otros pesos ω_k , de forma que:

1. El estimador basado en dichos pesos proporcione estimaciones perfectas para \mathbf{x} , es decir,

$$\sum_s \omega_k \mathbf{x}_k = T_{\mathbf{x}} = (T_{x_1}, \dots, T_{x_J}). \quad (5.1)$$

Estas son las ecuaciones de calibración (una por cada variable auxiliar), donde las estimaciones de los totales poblacionales de las variables auxiliares coinciden con los valores poblacionales conocidos. Así pues, los pesos ω_k reproducen exactamente los totales poblacionales conocidos para cada variable auxiliar del vector \mathbf{x} , es decir, proporciona estimaciones perfectas para las variables auxiliares. Esto debe ser así para garantizar que la muestra ponderada es representativa con respecto a las variables auxiliares utilizadas.

2. Esos nuevos pesos calibrados ω_k estén tan próximos como sea posible, respecto a una distancia dada, a los pesos originales d_k , con la expresión siguiente:

$$\min_{\omega_k} E \left\{ \sum_{k \in s} G(\omega_k, d_k) \right\}. \quad (5.2)$$

Así pues, ω_k y d_k tienen que ser tan cercanos como sea posible para que los estimadores resultantes sean insesgados o casi insesgados. Entre las posibles soluciones de estas ecuaciones de equilibrio se elegirá aquella en la que los nuevos pesos se diferencien lo menos posible de los antiguos. El problema de minimización puede ser resuelto usando el *método de los multiplicadores de Lagrange* (Deville y Särndal, 1992).

Usualmente la *distancia* elegida es la suma ponderada de cuadrados de las distancias, que viene dada por

$$\sum_{k \in s} \frac{(\omega_k - d_k)^2}{q_k d_k}, \quad (5.3)$$

donde q_k son constantes positivas. Entonces tenemos el siguiente problema:

$$\blacksquare \text{ Minimizar } \sum_{k \in s} \frac{(\omega_k - d_k)^2}{q_k d_k} \text{ sujeto a la condición } \sum_{k \in s} \omega_k \mathbf{x}_k = T_x.$$

Utilizando el método de los *multiplicadores de Lagrange* se obtienen los siguientes pesos calibrados:

$$\omega_k = d_k + d_k q_k \lambda \mathbf{x}'_k, \quad (5.4)$$

donde $\lambda = T_s^{-1}(T_x - \hat{T}_{xH})$, supuesto que la inversa de $T_s = \sum_{k \in s} d_k q_k \mathbf{x}_k \mathbf{x}'_k$ existe, y siendo \hat{T}_{xH} el estimador de *Horvitz-Thompson* para el vector de variables auxiliares \mathbf{x} .

El estimador calibrado obtenido, así construido viene dado por:

$$\hat{T}_{yreg} = \sum_{k \in s} \omega_k y_k = \hat{T}_{yH} + (T_x - \hat{T}_{xH}) \cdot \hat{B}_s, \quad (5.5)$$

siendo

$$\hat{B}_s = T_s^{-1} \cdot \sum_{k \in s} q_k \mathbf{x}_k y_k. \quad (5.6)$$

Dicho estimador, es el *estimador general de regresión* (Cassel y col. 1976; ver también [Subsección 2.3.3](#)).

La forma de \hat{T}_{yreg} dependerá tanto del diseño muestral como de las constantes q_k elegidas. Por ejemplo, si trabajamos con una única variable auxiliar $\mathbf{x} = x_1$, como q_k elegimos $q_k = \frac{1}{x_k}$ y bajo un [MAS](#) entonces

$$\hat{T}_{yreg} = \frac{\hat{T}_{yH}}{\hat{T}_{xH}} \cdot T_x,$$

que es el *estimador de razón* (Subsección 2.3.2).

En general, el estimador \hat{T}_{yreg} no es insesgado, pero como los pesos ω_k están próximos a d_k , es asintóticamente insesgado.

La varianza asintótica del estimador general de regresión viene dada aproximadamente por

$$AV(\hat{T}_{yreg}) = \sum_{k \in U} \sum_{l \in U} \Delta_{kl} (d_k E_k) (d_l E_l), \quad (5.7)$$

donde $\Delta_{kl} = \pi_{kl} - \pi_k \pi_l$ y $E_k = y_k - \mathbf{x}'_k B$, siendo

$$B = \left(\sum_{k \in U} \mathbf{x}_k \mathbf{x}'_k q_k \right)^{-1} \cdot \sum_{k \in U} q_k \mathbf{x}_k y_k.$$

Un estimador para esta varianza, viene dado por:

$$\widehat{V}_1(\hat{T}_{yreg}) = \sum_{k \in s} \sum_{l \in s} \Delta_{kl} d_{kl} (d_k e_k) (d_l e_l), \quad (5.8)$$

siendo $d_{kl} = \frac{1}{\pi_{kl}}$ y $e_k = y_k - \mathbf{x}'_k \widehat{B}_s$, donde \widehat{B}_s viene dado por [Ecuación 5.6](#) (para un estudio detallado de estos resultados puede consultarse [Särndal, Swensson y JH Wretman 1989](#)).

Otro posible estimador para la varianza de \hat{T}_{yreg} viene dado por

$$\widehat{V}_2(\hat{T}_{yreg}) = \sum_{k \in s} \sum_{l \in s} \Delta_{kl} d_{kl} (\omega_k e_k) (\omega_l e_l), \quad (5.9)$$

donde los pesos calibrados ω_k sustituyen a los pesos originales d_k . Este estimador suele resultar más adecuado que el anterior cuando trabajamos con un modelo de superpoblación apropiado ([Deville y Särndal, 1992](#)), por ejemplo

$$E_{\xi}(y_k) = \beta' \mathbf{x}_k \quad \text{y} \quad V_{\xi}(y_k) = \sigma_k^2.$$

A la vista de estos resultados, el caso ideal, aunque utópico, se daría cuando la relación entre las variable de interés y y las variables auxiliares es

$$y_k = \beta' \mathbf{x}_k \quad k \in U,$$

siendo β un vector de constantes. En tal caso la estimación coincide con el verdadero valor, es decir, $\hat{T}_{yreg} = T_y$ y la varianza de \hat{T}_{yreg} es nula.

5.3 MODIFICACIONES EN LA CONSTRUCCIÓN DE LOS ESTIMADORES DE CALIBRACIÓN

En esta sección, revisaremos otros estimadores calibrados, que se obtienen modificando la distancia elegida (Ecuación 5.2), así como modificando el proceso de construcción de un estimador de calibración. Dicho proceso se basa en:

1. Que los pesos calibrados den estimaciones perfectas para las variables auxiliares.
2. Minimizar una distancia.

Por tanto, las modificaciones que aquí revisaremos afectarán a una o a otra condición. Muchas de estas modificaciones harán uso de modelos de superpoblación para que den resultados deseables.

5.3.1 Estimadores de calibración para una familia de distancias

En la sección anterior se ha estudiado el estimador de calibración para una determinada distancia, obteniendo el estimador general de regresión, pero en otras ocasiones se pueden emplear otro tipo de distancias. En esta sección analizamos los estimadores de calibración obtenidos mediante una determinada familia de distancias. En concreto, la familia de distancias es la siguiente: vamos a considerar para cada $k \in s$, una función $G_k(\omega, d)$, tal que:

- Para todo valor fijo $d > 0$, $G_k(\omega, d)$ es no negativa, diferenciable con respecto a ω , estrictamente convexa, definida en un intervalo $D_k(d)$, que contiene a d y tal que $G_k(d, d) = 0$.
- $g_k(\omega, d) = \frac{\partial G_k(\omega, d)}{\partial \omega}$ es continua y es una aplicación biyectiva de $D_k(d)$ en un intervalo $Im_k(d)$.

Debido a estas condiciones tenemos que $g_k(d, d) = 0$ y además que $g_k(\omega, d)$ es una función estrictamente creciente de ω .

Una vez descritas estas condiciones, la distancia elegida es

$$\sum_{k \in s} G_k(\omega_k, d_k),$$

sujeto a la condición, ya conocida,

$$\sum_{k \in s} \omega_k \mathbf{x}_k = T_{\mathbf{x}}.$$

A partir de las condiciones impuestas, no podemos garantizar la existencia de solución, pero lo que sí se puede garantizar es que si existe es única, y además vendría dada por

$$\omega_k = d_k F_k(\mathbf{x}'_k \lambda),$$

donde λ es el vector de multiplicadores de Lagrange y $d_k F_k(\cdot)$ es la función inversa de $g_k(\cdot, d_k)$.

Debido a las condiciones impuestas, tenemos que $F_k(0) = 1$ y $F'_k(0) > 0$, de forma que notaremos por q_k a $F'_k(0)$.

En la mayoría de aplicaciones $g_k(\omega, d) = \frac{g(\omega/d)}{q_k}$ donde $g(\cdot)$ es una función real de variable real, derivable, estrictamente creciente, y tal que $g(1) = 0$ y $g'(1) = 1$. Si $F(u) = g^{-1}(u)$ entonces los pesos calibrados que se obtienen son:

$$\omega_k = d_k F(q_k \mathbf{x}'_k \lambda),$$

donde λ se determina mediante la siguiente ecuación:

$$\Phi_s(\lambda) = \sum_{k \in s} d_k \left(F_k(q_k \mathbf{x}'_k \lambda) - 1 \right) \mathbf{x}_k = T_x - \hat{T}_{xH},$$

que para poder resolver necesitaremos, en ocasiones, aplicar métodos tales como *Newton-Raphson*, etc. Ejemplos de este tipo de funciones, se pueden observar en [Tabla 5.1](#).

Tabla 5.1: Ejemplos de funciones $G_k(\omega_k, d_k)$.

Nombre distancia	$q_k G_k(\omega_k, d_k)$	$g(\omega_k/d_k) = q_k g_k(\omega_k, d_k)$	$F_k(u) = F(q_k u)$
Chi-cuadrado	$\frac{(\omega_k - d_k)^2}{2d_k}$	$\frac{\omega_k}{d_k} - 1$	$1 + q_k u$
Raking	$\omega_k \log\left(\frac{\omega_k}{d_k}\right) - \omega_k + d_k$	$\log\left(\frac{\omega_k}{d_k}\right)$	$\exp(q_k u)$
Hellinger	$2\left(\sqrt{\omega_k} - \sqrt{d_k}\right)^2$	$2\left(1 - \sqrt{\frac{d_k}{\omega_k}}\right)$	$\left(1 - \frac{q_k u}{2}\right)^{-2}$
Mínima entropía	$d_k \log\left(\frac{\omega_k}{d_k}\right) + \omega_k - d_k$	$1 - \left(\frac{\omega_k}{d_k}\right)^{-1}$	$\left(1 - q_k u\right)^{-1}$
Chi-cuadrado modificada	$\frac{(\omega_k - d_k)^2}{2\omega_k}$	$\frac{1 - \left(\frac{\omega_k}{d_k}\right)^{-2}}{2}$	$\left(1 - 2q_k u\right)^{-1/2}$

NOTA: Las funciones están normalizadas, así que $F'_k(0) = q_k$ en todos los casos

De esta manera, el *estimador de calibración* obtenido es

$$\hat{T}_{y\omega} = \sum_{k \in S} \omega_k y_k,$$

con $\omega_k = d_k F(q_k \mathbf{x}'_k \lambda)$. Estos estimadores también dan estimaciones perfectas cuando la relación entre la variable de estudio y y el vector de variables auxiliares \mathbf{x} , es lineal. Para un estudio más detallado de estos estimadores puede consultarse *Deville y Särndal (1992)*.

RELACIÓN CON EL ESTIMADOR GENERAL DE REGRESIÓN. A continuación, vamos a estudiar la relación existente entre esta familia de estimadores de calibración con el estimador general de regresión. Para ello es necesario imponer una serie de restricciones, para las que necesitamos considerar una sucesión de poblaciones finitas de tamaño N y de diseños muestrales, en las que seleccionaremos una muestra de tamaño n . El tamaño poblacional N tiende a infinito, cuando n tiende a infinito, y si \mathbf{x} es el vector de variables auxiliares, las condiciones que vamos a considerar son las siguientes:

1. $\lim_{n \rightarrow +\infty} N^{-1} T_{\mathbf{x}}$ existe.
2. $N^{-1}(\hat{T}_{\mathbf{x}H} - T_{\mathbf{x}}) \rightarrow 0$ con respecto a la probabilidad del diseño muestral.

3. $n^{1/2}N^{-1}(\widehat{T}_{xH} - T_x)$ converge en distribución a una normal multivariante $N(0, A)$, donde A es una matriz fija que describe el efecto asintótico del diseño muestral usado.

Bajo estas condiciones, se puede establecer el siguiente resultado. Para cualquier función $F_k(\cdot)$ perteneciente a la familia de distancias estudiadas (Tabla 5.1), obedeciendo las condiciones anteriores, tenemos que el estimador de calibración construido a partir de $F_k(\cdot)$, es asintóticamente equivalente al estimador general de regresión \widehat{T}_{yreg} , en el sentido de que

$$N^{-1}(\widehat{T}_{y\omega} - \widehat{T}_{yreg}) = O_p(n^{-1}).$$

Como consecuencia, los dos estimadores tienen la misma varianza asintótica dada por la Ecuación 5.7 (Deville y Särndal, 1992).

Por tanto, debido a este resultado podemos usar los estimadores de la Ecuación 5.8 y Ecuación 5.9, para estimar la varianza de cualquier estimador calibrado perteneciente a la familia de distancias estudiada, siempre y cuando estemos en las condiciones adecuadas.

5.3.2 Estimadores de calibración basados en una forma funcional

A continuación, el estudio se centrará en la modificación consistente en eliminar el proceso de minimización de la distancia entre d_k y ω_k e imponer que los pesos calibrados ω_k tengan una determinada forma funcional, en concreto:

$$\omega_k = d_k + q_k \lambda'_s \mathbf{z}_k,$$

donde q_k son constantes positivas, y \mathbf{z}_k es un vector, de forma que cumple las dos condiciones siguientes:

1. $\dim(\mathbf{z}_k) = J = \dim(\mathbf{x}_k)$
2. La matriz

$$\sum_{k \in s} q_k \mathbf{z}_k \mathbf{x}'_k$$

es no singular.

De esta manera, obtenemos los pesos

$$\omega_{kCALF} = d_k + (T_x - \widehat{T}_{xH})' R_k,$$

donde

$$R_k = \left(\sum_{k \in s} q_k \mathbf{z}_k \mathbf{x}'_k \right)^{-1} q_k \mathbf{z}_k,$$

y el estimador calibrado obtenido con estos pesos es

$$\hat{T}_{yCALF} = \hat{T}_{yH} + (T_x - \hat{T}_{xH})' \hat{Q},$$

donde

$$\hat{Q} = \left(\sum_{k \in S} q_k \mathbf{z}_k \mathbf{x}_k' \right)^{-1} \cdot \sum_{k \in S} q_k \mathbf{z}_k y_k.$$

RELACIÓN CON EL ESTIMADOR GENERAL DE REGRESIÓN. En primer lugar, podemos observar que el estimador general de regresión pertenece a esta nueva familia de estimadores calibrados, pues

$$\hat{T}_{yCALF} = \hat{T}_{yreg}$$

si $\mathbf{z}_k = \mathbf{x}_k$. Es decir, eligiendo el vector \mathbf{z}_k igual al vector de variables auxiliares obtenemos \hat{T}_{yreg} . Además, si existe una relación lineal entre el vector \mathbf{x} y la variable de estudio y , \hat{T}_{yCALF} da estimaciones perfectas.

Por otra parte, esta nueva familia de estimadores calibrados también, en general, son estimadores sesgados ya que

$$B[\hat{T}_{yCALF}] = -E[(T_x - \hat{T}_{xH})'(\hat{Q} - Q)] = O_p(n^{-1}),$$

donde

$$Q = \left(\sum_{k \in U} \pi_k q_k \mathbf{z}_k \mathbf{x}_k' \right)^{-1} \cdot \sum_{k \in U} \pi_k q_k \mathbf{z}_k y_k.$$

Este resultado puede consultarse en [VM Estevao y Särndal \(2006\)](#).

Respecto a las varianzas de estos nuevos estimadores comparadas con las del estimador general de regresión, exponemos el siguiente resultado, cuya demostración puede consultarse en [Särndal y VM Estevao \(2000\)](#).

Supongamos que para cada $k \in U$ existe un valor positivo q_k^* y un vector \mathbf{x}_k^* de variables auxiliares con total conocido $T_{\mathbf{x}^*} = \sum_{k \in U} \mathbf{x}_k^*$, tal que

$$\sum_{k \in U} \sum_{l \in U} \Delta_{kl} (d_k E_k^*) (d_l \mathbf{x}_l^*) = 0,$$

donde $E_k^* = y_k - (\mathbf{x}_k^*)'B^*$, con

$$B^* = \left(\sum_{k \in U} q_k^* \mathbf{x}_k^* (\mathbf{x}_k^*)' \right)^{-1} \left(\sum_{k \in U} q_k^* \mathbf{x}_k^* y_k \right).$$

Entonces el estimador \hat{T}_{yreg} tiene varianza asintótica mínima de entre todos los estimadores de la forma \hat{T}_{yCALF} .

En otras palabras, para este conjunto de variables auxiliares \mathbf{x}_k^* y para las constantes positivas q_k^* , tenemos que

$$AV(\hat{T}_{yCALF}) \geq AV(\hat{T}_{yreg})$$

para todos los estimadores \hat{T}_{yCALF} . Es decir, bajo ciertas condiciones, la varianza asintótica de \hat{T}_{yreg} es mínima dentro de la familia de estimadores \hat{T}_{yCALF} .

5.3.3 Estimadores de calibración cosméticos

Para introducir este tipo de estimadores calibrados, asumiremos el siguiente modelo de superpoblación:

$$y_k = \mathbf{x}_k' \beta + \epsilon_k \quad \text{con} \quad E_{\xi}[\epsilon_k] = 0, \quad E_{\xi}[\epsilon_k^2] = \sigma^2 a_k^2, \quad \text{y} \quad E_{\xi}[\epsilon_j \epsilon_k] = 0 \quad \forall k \neq j, \quad (5.10)$$

donde ϵ_k es un error aleatorio, σ^2 es un escalar desconocido y a_k^2 son constantes conocidas.

Una vez introducido el modelo de superpoblación, los *estimadores calibrados cosméticos*, surgen al intentar encontrar un estimador de β , β_{COSCAL} , de manera que la forma estándar y predictiva del estimador general de regresión sean numéricamente iguales, esto es

$$\hat{T}_{yCOSCAL} = \hat{T}_{yH} + (T_x - \hat{T}_{xH}) \hat{\beta}_{COSCAL} = \quad (5.11)$$

$$= \hat{T}_{ys} + (T_x - \hat{T}_{xs}) \hat{\beta}_{COSCAL}, \quad (5.12)$$

siendo \hat{T}_{ys} y \hat{T}_{xs} , los totales muestrales de y y de \mathbf{x} , respectivamente. Para conseguir este objetivo, es suficiente utilizar las técnicas de calibración ya revisadas, pero minimizando la distancia

$$\sum_{k \in s} \frac{(\omega_k - d_k)^2}{(d_k - 1)q_k}$$

sujeto a la condición $\sum_{k \in s} \omega_k \mathbf{x}_k = T_{\mathbf{x}}$ y siendo $q_k = \frac{1}{\mathbf{x}'_k \alpha}$, donde α es un vector de constantes. De esta forma obtenemos

$$\omega_k = d_k + (d_k - 1)q_k \mathbf{x}'_k \left(\sum_{k \in s} (d_k - 1)q_k \mathbf{x}_k \mathbf{x}'_k \right)^{-1} \left(T_{\mathbf{x}} - \hat{T}_{\mathbf{x}H} \right),$$

y el estimador de calibración, definido como $\sum_{k \in s} \omega_k y_k$, se reduce a la fórmula dada anteriormente para $\hat{T}_{yCOSCAL}$ (Ecuación 5.11).

Puesto que las formas Ecuación 5.11 y Ecuación 5.12 son iguales, existe otra fórmula alternativa para ω_k

$$\omega_k = d_k + (d_k - 1)q_k \mathbf{x}'_k \left(\sum_{k \in s} (d_k - 1)q_k \mathbf{x}_k \mathbf{x}'_k \right)^{-1} (T_{\mathbf{x}} - \hat{T}_{\mathbf{x}s}).$$

Si $y_k = \beta' \mathbf{x}_k$, siendo β un vector de constantes, entonces $\hat{T}_{yCOSCAL} = T_y$.

En relación a la varianza, puesto que este estimador $\hat{T}_{yCOSCAL}$ es el estimador general de regresión, tenemos el siguiente resultado:

Bajo el modelo de superpoblación (Ecuación 5.10), tenemos que la varianza con respecto al diseño muestral del estimador $\hat{T}_{yCOSCAL}$, viene dada por

$$V_p(\hat{T}_{yCOSCAL}) = V_p(\hat{T}_{yreg}) = \sum_{j=1}^N \sum_{k=1}^N (\pi_j \pi_k - \pi_{jk}) \cdot (\epsilon_j d_j - \epsilon_k d_k)^2.$$

En ocasiones, es más interesante estimar la varianza asintótica que la varianza con respecto al diseño, para ello presentamos el siguiente resultado:

Dado el modelo de superpoblación Ecuación 5.10, tenemos que la varianza asintótica de $\hat{T}_{yCOSCAL}$ viene dada por

$$AV[\hat{T}_{yCOSCAL} - T_y] = AV[\hat{T}_{yreg} - T_y] = E_{\xi} V_p(\hat{T}_{yreg}) = \sigma^2 \cdot \sum_{j=1}^N (d_j - 1) a_j^2.$$

Esta expresión, ya fue presentada por [Godambe \(1955\)](#), como la mínima varianza asintótica posible para cualquier estimador insesgado, con respecto al diseño, de T_y . Por tanto será más preferible, si el modelo de trabajo es útil, estimar la varianza asintótica, en lugar de la varianza con respecto al diseño.

Se sigue inmediatamente a partir de esta última expresión que un estimador de la varianza asintótica es

$$\hat{\sigma}^2 \cdot \sum_{k \in s} (d_j - 1) a_j^2,$$

donde $\hat{\sigma}^2$ puede ser un estimador de σ^2 obtenido a partir de una regresión basada en el uso de $\hat{\beta}_{\text{COSCAL}}$.

Por último, vamos a interesarnos por la varianza predictiva de $\hat{T}_{y\text{COSCAL}}$ y, para ello, presentamos el siguiente resultado:

Dado el modelo de superpoblación ([Ecuación 5.10](#)), tenemos que la varianza predictiva de $\hat{T}_{y\text{COSCAL}}$ viene dada por

$$E_{\xi} \left[\left(\hat{T}_{y\text{COSCAL}} - T_y \right)^2 \right] = \sigma^2 \left[\sum_{j \in s} \omega_j (\omega_j - 1) a_j^2 + \left(\sum_{j=1}^N a_j^2 - \sum_{j \in s} d_j a_j^2 \right) - \left(\sum_{j \in s} \omega_j a_j^2 - \sum_{k \in s} d_k a_k^2 \right) \right].$$

Asumiendo que a_j^2 son conocidos o pueden ser satisfactoriamente imputados, tanto la varianza predictiva como la varianza asintótica, pueden ser estimadas reemplazando en sus expresiones σ^2 por $\hat{\sigma}^2$.

Otra posible elección para estimar la varianza predictiva, sería

$$\hat{\sigma}^2 \sum_{k \in s} \omega_k (\omega_k - 1) a_k^2,$$

que se trata de un estimador simple y no requiere el conocimiento de los valores a_k^2 no muestrales.

Un estudio detallado de todos estos resultados puede consultarse en [Brewer \(1999\)](#).

5.3.4 Estimadores de calibración modelo-asistido

Los estimadores de calibración de este apartado surgen al sustituir en el proceso de calibración la restricción

$$\sum_{k \in S} \omega_k \mathbf{x}_k = T_{\mathbf{x}}$$

por otra restricción más adecuada.

Esta idea surge al plantearse cómo de efectivo es el uso que estamos haciendo de la información adicional a través de la anterior condición.

Ya comentamos en apartados anteriores, que usando la información adicional de esta manera, estamos implícitamente asumiendo un modelo lineal de regresión entre la variable de estudio y y las variables incluidas en el vector \mathbf{x} , en la población de estudio. Es decir, si asumimos una relación lineal $y_k = \mathbf{x}'_k \beta$ para todo $k \in U$, los estimadores de calibración hasta ahora estudiados coinciden y reproducen exactamente el total poblacional de la variable de estudio y .

Por tanto, los estimadores calibrados construidos hasta ahora, están motivados por un modelo lineal. Si otro tipo de curva es la que relaciona la variable de estudio y con las variables auxiliares que forman el vector \mathbf{x} , los estimadores de calibración vistos hasta el momento pueden resultar ineficaces.

De este modo, cuando trabajamos con otro tipo de relación entre \mathbf{x} e y , se hace necesario modificar la condición $\sum_{k \in S} \omega_k \mathbf{x}_k = T_{\mathbf{x}}$, y adaptarla a cada tipo de situación, para así poder usar la información auxiliar de la mejor manera posible.

A este tipo de estimadores calibrados, que se adaptan al modelo de trabajo, se les denomina estimadores calibrados modelo-asistidos.

A partir de ahora asumiremos que la relación entre las variable y y el vector \mathbf{x} puede ser descrita por el siguiente modelo de superpoblación:

$$E_{\xi}(y_k | \mathbf{x}_k) = \mu(\mathbf{x}_k, \theta) \quad V_{\xi}(y_k | \mathbf{x}_k) = v_k^2 \sigma^2, \quad (5.13)$$

con $k = 1, 2, \dots, N$ y donde $\theta = (\theta_0, \theta_1, \dots, \theta_J)'$ y σ^2 son parámetros poblacionales desconocidos, $\mu(\mathbf{x}, \theta)$ es una función conocida de \mathbf{x} y θ , v_k es una función conocida de \mathbf{x}_k o bien de $\mu_k = \mu(\mathbf{x}_k, \theta)$ y E_{ξ} y V_{ξ} denota

la esperanza y la varianza con respecto al modelo de superpoblación.

Asumiremos también que los pares $(y_1, \mathbf{x}_1); (y_2, \mathbf{x}_2); \dots; (y_N, \mathbf{x}_N)$ son mutuamente independientes.

Este modelo es bastante general, e incluye dos casos muy importantes:

1. El modelo de regresión lineal o no lineal

$$y_k = \mu(\mathbf{x}_k, \theta) + v_k \epsilon_k \quad k = 1, 2, \dots, N,$$

donde ϵ_k son variables aleatorias independientes e idénticamente distribuidas, con $E_{\xi}[\epsilon_k] = 0$ y $V_{\xi}(\epsilon_k) = \sigma^2$, y $v_k = v(\mathbf{x}_k)$ es una función conocida, estrictamente positiva que depende solamente de \mathbf{x}_k .

2. El modelo lineal generalizado.

$$g(\mu_k) = \mathbf{x}_k' \theta \quad V_{\xi}(y_k | \mathbf{x}_k) = v(\mu_k) \quad k = 1, 2, \dots, N,$$

donde $\mu_k = E_{\xi}(y_k | \mathbf{x}_k)$, $g(\cdot)$ es una función de enlace y $v(\cdot)$ es la función varianza.

Bajo el modelo [Ecuación 5.13](#), la información auxiliar debe ser usada a través de los valores $\mu(\mathbf{x}_k, \hat{\theta})$, con $k = 1, \dots, N$, donde $\hat{\theta}$ es una estimación de θ . Para ello, se define el estimador de calibración modelo-asistido de T_y como

$$\hat{T}_{yMC} = \sum_{k \in s} \omega_k y_k,$$

donde los pesos calibrados ω_k , minimizan una medida de distancia con respecto a d_k , sujeto a las nuevas condiciones

$$N^{-1} \cdot \sum_{k \in s} \omega_k = 1 \quad \sum_{k \in s} \omega_k \mu(\mathbf{x}_k, \hat{\theta}) = \sum_{k=1}^N \mu(\mathbf{x}_k, \hat{\theta}).$$

Minimizando la medida de distancia dada por

$$\sum_{k \in s} \frac{(\omega_k - d_k)^2}{d_k q_k} = \sum_{k \in s} \frac{d_k (\omega_k / d_k - 1)^2}{q_k}. \quad (5.14)$$

Sujeto a las condiciones anteriores, obtenemos el estimador

$$\hat{T}_{yMC} = \hat{T}_{yH} + \left(1 - \frac{1}{N} \sum_{k \in s} d_k\right) \left[\hat{Y} - \bar{\mu} \cdot \hat{B}_N\right] + \left(\sum_{k=1}^N \hat{\mu}_k - \sum_{k \in s} d_k \hat{\mu}_k\right) \hat{B}_N,$$

donde

$$\hat{B}_N = \frac{\sum_{k \in s} d_k q_k (\hat{\mu}_k - \bar{\mu})(y_k - \hat{Y})}{\sum_{k \in s} d_k q_k (\hat{\mu}_k - \bar{\mu})^2},$$

$$\hat{Y} = \frac{\sum_{k \in s} d_k q_k y_k}{\sum_{k \in s} d_k q_k}$$

y, por último,

$$\bar{\mu} = \frac{\sum_{k \in s} d_k q_k \hat{\mu}_k}{\sum_{k \in s} d_k q_k}.$$

Si la restricción $N^{-1} \sum_{k \in s} \omega_k = 1$ es eliminada, la ecuación de calibración

$$\sum_{k \in s} \omega_k \mu(\mathbf{x}_k, \hat{\theta}) = \sum_{k=1}^N \mu(\mathbf{x}_k, \hat{\theta})$$

produce el estimador

$$\hat{T}_{yMC}^* = \hat{T}_{yH} + \left(\sum_{k=1}^N \hat{\mu}_k - \sum_{k \in s} d_k \hat{\mu}_k \right) \hat{B}_N^*,$$

donde

$$\hat{B}_N^* = \frac{\sum_{k \in s} d_k q_k \hat{\mu}_k y_k}{\sum_{k \in s} d_k q_k \hat{\mu}_k^2}.$$

PROPIEDADES. Las propiedades más importantes de estos dos nuevos estimadores, son resumidas en el siguiente teorema, pero para poder establecerlo, son necesarias las siguientes condiciones:

1. $\hat{\theta} = \theta_N + O_p(n^{-1/2})$ y $\theta_N \rightarrow \theta$, donde θ_N , es la estimación de θ , realizada con todas las unidades de la población.
2. Para cada \mathbf{x}_k , $\frac{\partial \mu(\mathbf{x}_k, t)}{\partial t}$ es continua en t y

$$\left| \frac{\partial \mu(\mathbf{x}_k, t)}{\partial t} \right| \leq h(\mathbf{x}_k, \theta)$$

para t en un entorno de θ , y $N^{-1} \sum_{k=1}^N h(\mathbf{x}_k, \theta) = O_p(1)$.

3. Los pesos básicos muestrales, $d_k = \frac{1}{\pi_k}$, hacen que los estimadores de *Horvitz-Thompson* para ciertos totales poblacionales estén asintóticamente distribuidos normalmente.
4. Para cada \mathbf{x}_k , $\frac{\partial^2 \mu(\mathbf{x}_k, t)}{\partial t \partial t'}$ es continua en t y

$$\max_{j,k} \left| \frac{\partial^2 \mu(\mathbf{x}_i, t)}{\partial t_j \partial t_k} \right| \leq k(\mathbf{x}_i, \theta)$$

para t en un entorno de θ , y $N^{-1} \sum_{i=1}^N k(\mathbf{x}_i, \theta) = O_p(1)$.

Con estas condiciones, podemos establecer el siguiente resultado:

1. Asumiendo el modelo de superpoblación [Ecuación 5.13](#), y bajo las condiciones (1) \sim (3), tanto \hat{T}_{yMC} como \hat{T}_{yMC}^* son iguales a $\hat{T}_{yH} + O_p(n^{-1/2})$ y son de este modo, asintóticamente insesgados para T_y , con respecto al diseño, sin tener en cuenta si el modelo es correcto o no.
Estos estimadores son también aproximadamente insesgados bajo el modelo, supuesto la condición (1).
2. Bajo las condiciones (1) \sim (4) dadas previamente, la varianza asintótica, con respecto al diseño, de \hat{T}_{yMC} viene dada por

$$AV(\hat{T}_{yMC}) = \sum_{i=1}^N \sum_{j>i}^N (\pi_i \pi_j - \pi_{ij}) \left(\frac{U_i}{\pi_i} - \frac{U_j}{\pi_j} \right)^2,$$

donde π_{ij} son las probabilidades de inclusión de segundo orden, $U_i = y_i - A_N - \mu_i B_N$, $\mu_i = \mu(\mathbf{x}_i, \theta_N)$, con

$$B_N = \frac{\sum_{k=1}^N q_k (\mu_k - \bar{\mu}_N) (y_k - \hat{Y})}{\sum_{k=1}^N q_k (\mu_k - \bar{\mu}_N)^2},$$

$$A_N = \hat{Y}_N - \bar{\mu}_N \cdot B_N$$

$$\text{y donde } \bar{\mu}_N = \sum_{k=1}^N \mu_k \quad e \quad \hat{Y}_N = \sum_{k=1}^N y_k.$$

Un estimador de $AV(\widehat{T}_{yMC})$ viene dado por

$$\widehat{AV}(\widehat{T}_{yMC}) = \sum_{i < j}^n \frac{\pi_i \pi_j - \pi_{ij}}{\pi_{ij}} \left(\frac{u_i}{\pi_i} - \frac{u_j}{\pi_j} \right)^2,$$

donde $u_i = y_i - \widehat{A}_N - \widehat{\mu}_i \widehat{B}_N$.

Un resultado similar tenemos para \widehat{T}_{yMC}^* , si eliminamos \widehat{A}_N y reemplazamos \widehat{B}_N por \widehat{B}_N^* .

3. Si $q_k = \frac{1}{v_k^2}$, en la expresión [Ecuación 5.14](#), entonces \widehat{T}_{yMC} y \widehat{T}_{yMC}^* se reducen al estimador convencional de calibración (o al estimador general de regresión; [Deville y Särndal 1992](#)), cuando el modelo de trabajo usado es el lineal, donde $\mu(\mathbf{x}_k, \theta) = \theta_0 + \theta_1 \mathbf{x}_{k1} + \dots + \theta_J \mathbf{x}_{kJ}$.

La demostración de este resultado, se puede consultar en [Changbao Wu y Randy R Sitter \(2001\)](#).

De este modo, tanto \widehat{T}_{yMC} como \widehat{T}_{yMC}^* son modelo-asistidos, en el sentido de que pueden manejar tanto modelos lineales como no lineales, esto es, ambos son consistentes respecto al diseño, sin tener en cuenta el modelo y particularmente eficientes si el modelo es correcto. También, en el caso de un modelo sin error (esto es, $y_k = \mu_k$), tenemos que

$$\widehat{T}_{yMC} = \widehat{T}_{yMC}^* = T_y$$

5.3.5 Estimadores de calibración para la función de distribución y cuantiles

La teoría de muestreo en poblaciones finitas se ha centrado clásicamente en la estimación de parámetros poblacionales de tipo puntual como totales, medias, proporciones y varianzas, pero existen una serie de parámetros de tipo funcional que pueden proporcionarnos información relevante acerca del comportamiento global de la población.

En particular y básicamente en las dos últimas décadas, se han desarrollado una gran cantidad de investigaciones acerca de la estimación de la función de distribución poblacional asociada a una variable de estudio numérica, definida sobre la población.

El problema de la estimación de la función de distribución es importante tanto por el interés de la propia función de distribución como por su relación con otros parámetros de tipo no funcional, pues a través de la función de distribución se pueden estimar la mediana, los cuantiles, el índice de Gini y otros indicadores como los relacionados con determinantes sociales de la salud ([A Cabrera-León y col., 2016](#)).

Si consideramos una población finita $U = \{1, 2, \dots, N\}$, de tamaño N , sobre la cual está definida una variable de estudio y , y un vector de dimensión J , de variables auxiliares $\mathbf{x} = (x_1, x_2, \dots, x_J)$. La función de distribución de la variable y , viene dada por:

$$F_y(t) = \frac{1}{N} \sum_{k \in U} \Delta(t - y_k)$$

donde $\Delta(t - y_k)$ denota la siguiente función:

$$\Delta(t - y_k) = \begin{cases} 0 & \text{si } t < y_k \\ 1 & \text{si } t \geq y_k \end{cases}$$

Supongamos que una muestra s , es extraída de la población U , empleando un diseño muestral d , con probabilidades de inclusión dadas en la matriz de diseño $\Pi = \{\pi_{kj}\}_{1 \leq k, j \leq N}$, de manera que la variable y , es conocida sólo en la muestra s , mientras que el vector de variables auxiliares es conocido en toda la población. Entonces podemos estimar $F_y(t)$, mediante el estimador de *Horvitz-Thompson*, dado por:

$$\hat{F}_{yH}(t) = \frac{1}{N} \sum_{k \in s} \frac{\Delta(t - y_k)}{\pi_k} = \frac{1}{N} \sum_{k \in s} d_k \Delta(t - y_k)$$

donde $d_k = 1/\pi_k$.

[Harms y Duchesne \(2006\)](#); [Rueda, S. Martínez y col. \(2007\)](#) usan diferentes métodos para implementar la técnica de calibración a la estimación de la función de distribución y de cuantiles. El método propuesto por [Rueda, S. Martínez y col. \(2007\)](#) es más sencillo computacionalmente y se basa en obtener un estimador de calibración minimizando la expresión:

$$\frac{1}{2} \sum_{k \in s} \frac{(\omega_k - d_k)^2}{d_k q_k}$$

donde q_k son constantes positivas, y sujeto a la condición:

$$\frac{1}{N} \sum_{k \in s} \omega_k \Delta(t_0 - g_k) = F_g(t_0)$$

siendo $g = \hat{\beta}' \mathbf{x}$,

$$\hat{\beta} = \left(\sum_{k \in s} d_k q_k \mathbf{x}'_k \mathbf{x}_k \right)^{-1} \cdot \sum_{k \in s} \mathbf{x}_k y_k$$

un estimador de los coeficientes de regresión lineal entre la variable y de estudio y el vector de variables auxiliares \mathbf{x} , $F_g(t)$ la función de distribución de la variable g , y t_0 un valor elegido arbitrariamente.

Dicho estimador puede ser escrito de la forma:

$$\hat{F}_{yc}(t) = \hat{F}_{yH}(t) + \frac{(F_g(t_0) - \hat{F}_{gH}(t_0))}{\sum_{k \in s} d_k q_k \Delta(t_0 - g_k)} \sum_{k \in s} d_k q_k \Delta(t_0 - g_k) \Delta(t - y_k)$$

que es el estimador general de regresión para $\hat{F}_y(t)$ empleando como variable auxiliar g_k .

Los autores demuestran que el estimador es asintóticamente normal e insesgado para la función de distribución poblacional, y extienden el estimador al caso de que existan varias variables auxiliares.

La precisión del estimador depende del valor t_0 elegido para la calibración. En [S. Martínez, Rueda, Antonio Arcos y H. Martínez \(2010\)](#) se determina el punto óptimo, t_0 , en sentido de mínima varianza.

Los pesos de calibración pueden también obtenerse minimizando la distancia a los pesos originales pero imponiendo la calibración de la función de distribución F_g en P puntos fijados t_1, t_2, \dots, t_P . En [S. Martínez, Rueda, Antonio Arcos, H. Martínez y J. Muñoz \(2012\)](#); [S. Martínez, Rueda, H. Martínez y col. \(2015\)](#) se estudia el problema de la determinación del vector óptimo para la calibración.

[Kovacevic \(1997\)](#); [S. Martínez, Rueda, Antonio Arcos, H. Martínez y J. Muñoz \(2012\)](#) consideran nuevos estimadores de cuantiles a partir de los estimadores de calibración de la función de distribución y determinan las propiedades de dichos estimadores.

5.3.6 Estimadores de calibración en dominios

En las encuestas no sólo necesitamos la estimación del total de la población, sino también en otros dominios o subpoblaciones de interés.

Como hemos visto en los apartados anteriores, para reducir la varianza del estimador de calibración, debemos seleccionar aquel vector de variables auxiliares que mejor explique la variación de las variables de estudio principales. Además, relacionado con dominios, debemos tener también en cuenta que ese vector sea el que identifique los dominios más importantes ([Lundström y Särndal, 2001](#)).

Cuando se cumple el primer principio, el sesgo de cobertura es reducido en las estimaciones para las variables de estudio principales. Entonces la varianza de esas estimaciones serán igualmente reducidas.

Con respecto al segundo principio, su efecto al cumplirse se da fundamentalmente en una reducción de la varianza para las estimaciones en los dominios.

Si la información auxiliar esta formada por vectores con $\sum_{k \in U} \mathbf{x}_k$ conocido, sus pesos estarán dados por un conjunto de pesos calibrados descritos en Sección 5.2. Estos han sido utilizados para producir el estimador del total de T en la población. Para el total del dominio T_d , mantenemos los mismos pesos calibrados y cambiamos solamente la variable de estudio y por y_d . El estimador de calibración resultante para el total del dominio sería

$$\hat{T}_{y_d\omega} = \sum_{k \in s} \omega_k y_{dk}.$$

En algunas aplicaciones los dominios de interés $U_1, \dots, U_d, \dots, U_D$ forman una partición de U , como cuando los dominios son las provincias que componen una región D . Los estimadores de los D dominios $\hat{T}_1, \dots, \hat{T}_{y_d\omega}, \dots, \hat{T}_D$, tiene la propiedad atrayente de que su suma da la estimación de calibración para el conjunto de la población. Esta propiedad viene de:

$$\sum_{d=1}^D \hat{T}_{y_d\omega} = \sum_{d=1}^D \sum_{k \in s} \omega_k y_{dk} = \sum_{k \in s} \omega_k \sum_{d=1}^D y_{dk} = \sum_{k \in s} \omega_k y_k = \hat{T}_y.$$

Así pues, a la hora de obtener las estimaciones por dominios, utilizaríamos variables indicadoras del dominio correspondiente. Sea I_k^d , la función indicadora del dominio ($=1$ si el elemento pertenece al dominio k y 0 en caso contrario). Entonces, la variable auxiliar a considerar para la calibración en dominios estaría formada por $(I_k^1, \dots, I_k^D, xI_k^1, \dots, xI_k^D)$. Las restricciones se engloban utilizando una matriz de información auxiliar con las variables indicadoras de área junto con la variable auxiliar observada en cada área, es decir,

$$\sum_s \omega_k x_k = (N_1, \dots, N_D, X_1, \dots, X_D) y, \text{ por tanto, } \sum_s \omega_k = N y \sum_s \omega_k x = X.$$

Una vez obtenidos los pesos calibrados conjuntos, ω_k , el estimador para el total del dominio d sería

$$\hat{T}_{y_d\omega} = \sum_{k \in s} \omega_k y_k I_k^d.$$

VARIABLES AUXILIARES NO CONOCIDAS EN DOMINIOS. En ocasiones, la variable auxiliar observada en la muestra en un dominio es constante o tiene muy poca variación. También puede ocurrir que el total de la variable auxiliar se conozca solo para el total poblacional y no para cada dominio. Para esos casos, usaríamos una restricción en la calibración global, manteniendo las restricciones sobre los tamaños de los dominios, de forma que tendríamos $D+1$ restricciones, es decir,

$$\sum_s \omega_k x_k = (N_1, \dots, N_D, X) \text{ y, de esta forma, } \sum_s \omega_k = N \text{ y } \sum_s \omega_k x = X.$$

Con esos pesos comunes, aplicaríamos el método anterior.

CÁLCULO DE LAS VARIANZAS ESTIMADAS. Para obtener las estimaciones de las varianzas en los dominios, y por tanto de los errores, usaríamos el método de los residuos restringidos. Estos usan los pesos de calibración comunes y los coeficientes de calibración restringidos por dominio. Es decir, la predicción de los residuos tiene un peso común de calibración pero se restringe a los valores observados de las variables involucradas en el dominio correspondiente; es la varianza estimada de los residuos en el dominio, de forma que, $e_{kd} = y_{kd} - \beta_d x_k$.

Finalmente, como hemos comentado al principio, en la calibración por dominios es importante identificar información auxiliar lo más cercana posible a la identificación de los dominios. Esto reducirá el tamaño absoluto de los residuos, que resultará en una menor varianza. Si no es así, su efecto será leve, especialmente para los dominios más pequeños.

5.4 OTROS TRABAJOS SOBRE CALIBRACIÓN

La idea de la calibración ha tenido una gran aceptación entre los teóricos de muestreo en poblaciones finitas y han surgido así una gran cantidad de trabajos que utilizan el método de calibración bajo distintos tipos de diseños muestrales, así como para la estimación de otros parámetros distintos al total poblacional.

Algunos de estos trabajos a destacar son los siguientes:

[Dupont \(1995\)](#) considera estimadores de calibración de un total cuando la información auxiliar es obtenida en diferentes fases mediante un procedimiento de muestreo bifásico.

[Firth y Bennett \(1998\)](#) utilizaron modelos de regresión no lineales en la estimación de calibración.

[Welsh y Ronchetti \(1998\)](#) proponen estimadores calibrados bajo la suposición de que la muestra contiene outliers representativos.

Deville (2005); Harms (2003); Lundström y Särndal (1999); Särndal y Lundström (2005) aplican la técnica de calibración para el tratamiento de la no respuesta, reduciendo tanto el error muestral como el sesgo de no respuesta. Igualmente discuten el uso de la imputación como complemento a la ponderación por calibración.

Théberge (1999) considera estimadores de calibración para estimar cualquier parámetro lineal dando una solución al caso de que no exista solución al problema de calibración mediante inversas de Moore-Penrose.

S Singh y col. (1999) investigan nuevas técnicas de calibración para estimadores de la varianza de medias muestrales, estimadores de razón y regresión bajo diferentes esquemas de muestreo. S Singh (2001) también propone estimadores de la varianza poblacional.

Särndal y Lundström (2008, 2009), describen cómo evaluar y elegir el vector auxiliar que más reduce tanto el sesgo como el error muestral.

Park y Fuller (2009) consideraron el modelo de superpoblación de regresión con componentes aleatorios.

Montanari y Ranalli (2009) propusieron redes neuronales como técnica de suavizamiento multivariante para la estimación de calibración modelo-asistida no paramétrica (ver Sección 4.3).

El trabajo de Lee y Valliant (2009) sobre encuestas panel Web voluntarias está teniendo igualmente un gran impacto en la comunidad científica, sirviendo de referencia para otros muchos trabajos (S. Chen y Kim, 2014; Guzi y Garcia, 2015; Hill y Shaw, 2013; Schober y col., 2016). Combinan propensity score y calibración para reducir el sesgo de selección debido a la no aleatorización de la muestra y para corregir el sesgo de cobertura, respectivamente. Con los pesos obtenidos comparan 3 estimadores de la varianza, decantándose por el del método de replicación de muestras (jackknife). Por ejemplo, uno de los artículos recientemente publicado en el ámbito de la salud (posterior al nuestro de A. Cabrera-León y col. 2015), es el de Legleye y col. (2015), donde comparan mediante estimadores calibrados dos encuestas sobre salud reproductiva, una mediante un panel Web voluntario y la otra mediante una encuesta telefónica.

RJ Little (2012) basa las técnicas de calibración aplicando inferencia bayesiana (Fienberg, 2011; Gelman y col., 2013; RJ Little, 2003). Esta aproximación requiere la especificación de un modelo probabilístico completo (distribuciones a priori sobre parámetros), basando las inferencias en la distribución predictiva a posteriori de los valores no muestreados de la variable de interés. Entre otras ventajas, RJ Little (2012) identifica que se trata de una aproximación unificada para muestras grandes y pequeñas, errores de respuesta y de no respuesta (RJ Little y Rubin, 2016), para la fusión de datos, combinar fuentes con infor-

mación probabilística y no probabilística y para, en definitiva, el tratamiento de Big Data como argumenta en su artículo [RJ Little \(2015\)](#). Los avances computacionales para la aplicación factible de los métodos bayesianos (p.ej. rutinas de integración numérica o las técnicas de simulación), están dando a la calibración bayesiana un rol cada vez más importante en la investigación por muestreo.

Esos y otros trabajos realizados en los últimos años ([Barranco-Chamorro, Jiménez-Gamero, Mayor Gallego y col., 2015](#); [Barranco-Chamorro, Jiménez-Gamero, Moreno-Rebollo y col., 2012](#); [Dykes y col., 2015](#); [S. Martínez, Antonio Arcos y col., 2015](#); [S Singh, 2013](#)) pueden darnos una idea de lo rápido que se ha extendido esta nueva forma de utilizar la información auxiliar, aplicándose a muy variados problemas de estimación. Si bien en su comienzo los estimadores calibrados tenían el inconveniente de que su utilización sólo era sencilla en un caso muy concreto de modelos de superpoblación y para una cierta métrica, los últimos trabajos han permitido extender la técnica para modelos de superpoblación muy generales, parámetros más amplios y una amplia gama de métricas, resolviendo los problemas de no existencia de solución de la ecuación de calibración. Además permiten englobar como casos particulares a los estimadores más usuales.

Por último, la implementación de la calibración ha dado lugar igualmente a numerosos desarrollos informáticos, como el paquete estadístico `FRAMES2` (ver [Sección 7.3](#)), realizado en el programa R por [Antonio Arcos, Molina y col. \(2015\)](#), que incluye los principales estimadores en encuestas con marcos duales. Dado que una de las aplicaciones mostradas en esta tesis está relacionada con calibración para marcos imperfectos ([Capítulo 8](#)), se desarrolla esta metodología en el siguiente [Capítulo 6](#).

5.5 VENTAJAS Y COMPARACIÓN DE LA CALIBRACIÓN

De manera resumida, como ventajas principales de la calibración de estimadores tendríamos:

- Se trata de una metodología de gran importancia en la producción reciente de estadísticas a gran escala y de publicaciones científicas.
- Permite utilizar la información auxiliar existente en cualquier etapa o nivel de un muestreo por conglomerados o polietápico.
- Puede ofrecer un sistema de ponderación único, válido para todas las variables de estudio.
- Muchos de los estimadores indirectos no precisan de un modelo.

- Ofrece estimaciones más precisas y consistentes, es decir, de menor varianza (asintótica) sin importar cómo se elija la muestra y, de manera especial, en subpoblaciones pequeñas, en encuestas combinadas o en encuestas de hogares.
- Permite reducir el sesgo de no respuesta o por valores extremos y el de no cobertura.
- Y, por último, es una metodología transparente y natural, en el sentido de que las estimaciones resultantes son fáciles de interpretar y motivan, por lo tanto, al uso de la calibración.

En cuanto a la comparación de la calibración con otros métodos de estimación indirectos, podemos destacar lo siguiente:

- Los estimadores indirectos de razón, regresión, post-estratificación y diferencia, son casos particulares de la calibración bajo ciertas distancias y ecuaciones de equilibrio.
- Con respecto a la verosimilitud empírica:
 - El proceso de maximizar la verosimilitud, es equivalente a minimizar la distancia de Kullback-Leibler, pudiéndose ver, por tanto, la verosimilitud empírica como un caso particular de la calibración con esa distancia.
 - Por otro lado, el procedimiento de verosimilitud empírica está menos implementado en Software porque el proceso para maximizar la verosimilitud es menos eficiente que el de calibración, de ahí que las agencias estadísticas suelen utilizar calibración en vez verosimilitud empírica.
- En cuanto a la regresión no paramétrica:
 - Está desarrollada para variables (de estudio) cuantitativas.
 - Necesita que la información auxiliar sea completa, es decir, conocer todos sus valores en la población de estudio.
 - Además, requiere mucho más tiempo de computación y es difícil seleccionar el ancho de banda.

CALIBRACIÓN PARA MARCOS IMPERFECTOS

En el [Sección 1.4](#) se introdujeron algunos de los errores ajenos al muestreo más frecuentes en el desarrollo de una encuesta poblacional, entre los que se encuentra el error de cobertura. Así pues, la primera suposición que se asume al hacer inferencia basada en una muestra es que ésta se selecciona de un único marco perfecto que cubre a todos los individuos de la población objeto de estudio. Sin embargo, en la práctica apenas ocurre eso, siendo el marco y la población objeto de estudio diferentes. Por ejemplo, las encuestas por teléfonos fijos o móviles son un caso muy claro de encuestas con marcos imperfectos. Pero sobre todo en las encuestas Web este problema es mucho más grave (el 25,3 % de la población española entre 16 y 74 años no son usuarios frecuentes de Internet; [INE 2015a](#)). Esto produce un sesgo de cobertura que se agrava si las características de la población cubierta y no cubierta por el marco difieren. Así pues, está claro que variables como la edad, el nivel educacional y económico de la población que no tiene Web o teléfono móvil no son iguales al del resto de la población y, si bien estas diferencias sociodemográficas pueden ser ajustadas mediante ciertos procedimientos, existen también diferencias entre otras variables como estilo de vida, preferencias, comportamientos, que no pueden ser controladas. Esto hace que los sesgos de cobertura persistan en las inferencias realizadas para la población.

En este capítulo vamos a estudiar los problemas de estimación derivados de la no cobertura de la población en el marco y veremos diversos procedimientos para reducir estos problemas.

6.1 INTRODUCCIÓN

La mayor imperfección del marco muestral poblacional ocasiona un incremento del error de cobertura, con consecuencias en un incremento del sesgo de la estimación y en la pérdida de efectividad a la hora de seleccionar la muestra mediante estratificación o para la post-estratificación.

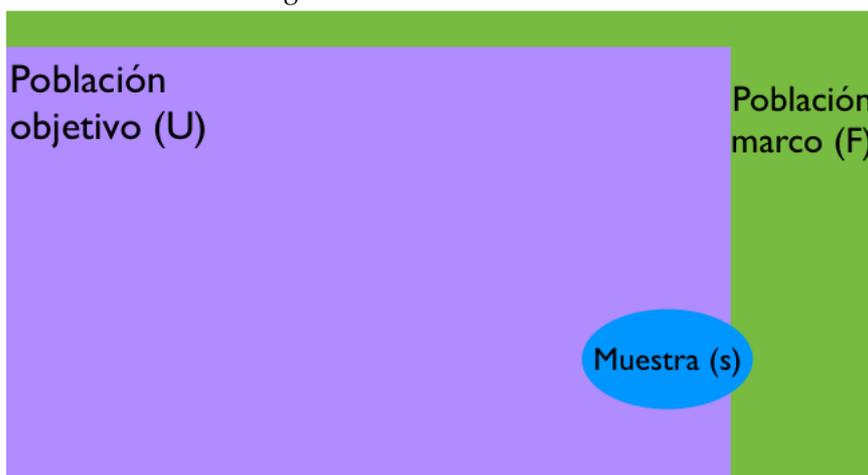
Los problemas más frecuentes derivados de marcos imperfectos son:

- La sobrecobertura (o exceso de cobertura): ocurre cuando una unidad de muestreo se encuentra en la población marco pero no en la población de estudio (unidades extrañas), bien porque se desea estudiar en cualquier caso o bien porque no es una unidad del

colectivo que se desea seleccionar para ser parte de la muestra. En este caso, la población de estudio se convierte en una subpoblación del marco muestral (Figura 6.1). Las unidades que han dejado de existir son fuentes frecuentes de sobrecobertura (conocidas como muertes). La sobrecobertura puede ser evitada si se identifican las unidades de la muestra a la que pertenece. Algunos ejemplos serían:

- El marco muestral está formado por viviendas ocupadas y vacías. En la población objetivo solo hay viviendas ocupadas. Al seleccionar una muestra puede darse el caso de tener unidades extrañas (vacías).
- La población de estudio está formada por mujeres entre 30 y 49 años. Sin embargo, el marco muestral está compuesto por mujeres de cualquier edad.

Figura 6.1: Error de sobrecobertura

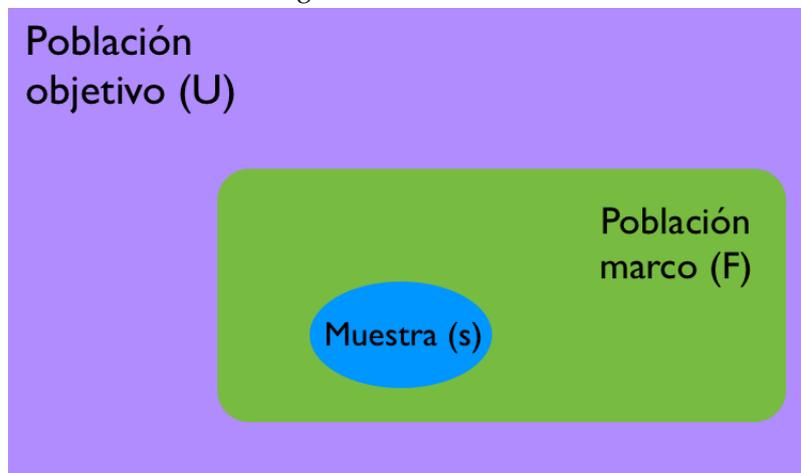


- La duplicación (o listados dobles): se da cuando las unidades muestrales son registradas en el marco muestral más de una vez. Nuevamente, según los ejemplos anteriores:
 - Una vivienda con dos personas se registra en el censo de 2001 (marco muestral) como ocupada. En los siguientes años, estas personas abandonan la vivienda que es nuevamente ocupada por otras 3 personas diferentes. Esta vivienda se registra en el siguiente censo de 2011 como ocupada, sin eliminar la vivienda registrada en el censo de 2001.
 - Un mujer perteneciente a la población de estudio es registrada en el marco muestral en la vivienda donde reside con su

madre los fines de semana y también en la vivienda donde reside con sus hijos entre semana.

- La subcobertura (o defecto de cobertura): se presenta cuando existen unidades que están en la población objetivo pero no en la población marco (Figura 6.2). Al contrario que con la sobrecobertura, el marco muestral se convierte en un subconjunto de la población. Una parte importante de la subcobertura esta formada por las unidades nuevas de la población objetivo (conocidas como nacimientos), que no están en el marco debido por ejemplo a no estar actualizado. Estas unidades en subcobertura tienen probabilidad cero de ser seleccionadas en cualquier muestra de la población marco. Siguiendo los ejemplos anteriores:
 - Se han construido nuevas viviendas en una ciudad que son parte de la población de estudio y que la última edición del censo (marco muestral) no las incluyó ya que este fue anterior.
 - La población de estudio está formada por mujeres entre 30 y 49 años residentes en España, seleccionando una muestra aleatoria de aquellas residentes solo en un municipio.

Figura 6.2: Error de subcobertura



A continuación se describe el problema del sesgo que produce el error de subcobertura.

Sean:

- $\mathcal{U} = \{1, \dots, k, \dots, N\}$ la población de estudio constituida por N unidades.

- $F = \{1, \dots, l, \dots, L\}$ un marco poblacional de \mathcal{U} constituido por L elementos e imperfecto, es decir, que no cubre a todas las unidades de la población \mathcal{U} .
- $s = \{1, \dots, n\}$ una muestra aleatoria de tamaño n seleccionada de F .
- $C = \{1, \dots, c\}$ los elementos de la población de estudio que no forman parte de F con tamaño $c = N - L$ (Figura 6.2).

El diseño muestral $d(s)$ definido en F induce probabilidades de primer orden $\pi_k > 0$ para todo $k \in s$ y pesos muestrales $d_k = 1/\pi_k$. Así pues, el total poblacional de una variable de estudio y es $T_y = \sum_{k=1}^N y_k$ y el estimador de *Horvitz y D. Thompson (1953)* de y bajo el diseño muestral elegido sería:

$$\hat{T}_{yH} = \sum_{k \in s} \frac{1}{\pi_k} y_k = \sum_{k \in s} d_k y_k.$$

El sesgo de \hat{T}_{yH} vendría dado por

$$B(\hat{T}_{yH}) = E[\hat{T}_{yH}] - E[T_y] = T_{yF} - T_{yU},$$

por lo que el estimador \hat{T}_{yH} es sesgado en cuanto que $F \subset U$. A continuación se muestran diversos procedimientos para reducir este sesgo.

6.2 UN MARCO IMPERFECTO

Una vía para tratar los problemas derivados de la falta de cobertura de un marco imperfecto es asumir un diseño muestral bifásico, donde el marco F es una muestra de una hipotética población U y la muestra s se extrae de F en una segunda fase de manera independiente.

Supongamos que asociada a la población U tenemos $\mathbf{x} = (x_1, \dots, x_j)$, un vector de variables auxiliares de forma que es perfectamente conocido para todos los elementos de la población U (T_x es conocido). Entonces la probabilidad de que una unidad $k \in U$ pertenezca al marco F es modelizada por

$$p_k = \frac{1}{f(h_k \gamma)},$$

donde γ es desconocida, h_k es un vector con la misma dimensión que \mathbf{x}_k y la matriz $\sum_{k \in s} d_k f(h_k \gamma) h_k' \mathbf{x}_k / N$ es invertible. La función se asume

monótona y dos veces diferenciable con derivadas segundas acotadas. Su forma funcional es conocida aunque el parámetro γ no.

Por otra parte, consideramos la metodología de la calibración basada en una forma funcional (Subsección 5.3.2), escogiendo la forma funcional más general. Siguiendo esta idea eliminamos el proceso de minimización de la distancia entre d_k y ω_k e imponemos que los pesos calibrados ω_k tengan una determinada forma funcional, en concreto:

$$\omega_k = d_k f(h_k g)$$

sujeto a la condición, ya conocida,

$$\sum_{k \in S} \omega_k \mathbf{x}_k = T_{\mathbf{x}}.$$

Usando el método iterativo descrito en Phillip S Kott (2009) para hallar \mathbf{g} obtenemos un estimador de calibración:

$$T_{y_{cal}} = \sum_{k \in S} d_k f(h_k g) y_k.$$

Este estimador tendrá buenas propiedades bajo el modelo de predicción lineal, $y_k = \beta \mathbf{x}'_k + \epsilon_k$ cuando cada h_k es una función de \mathbf{x}_g ($g \in U$) y $E(\epsilon_k | \{\mathbf{x}_g, I_g; g \in U\}) = 0$ para todo $k \in U$, siendo I_g una variable indicadora de que la unidad está tanto en la muestra como en el marco.

La condición $\sum_{k \in S} \omega_k \mathbf{x}_k = T_{\mathbf{x}}$, junto con el teorema del valor medio, proporcionan el siguiente resultado:

$$T_{\mathbf{x}} = \sum_{k \in S} d_k f(h_k \gamma) \mathbf{x}_k + \sum_{k \in S} d_k f'(\theta_k) (h_k (g - \gamma)) \mathbf{x}_k = O_p\left(\frac{N}{n^{1/2}}\right),$$

para algún valor θ_k entre $h_k g$ y $h_k \gamma$. Entonces si $\sum_{k \in S} d_k f(h_k \gamma) h'_k \mathbf{x}_k / N$ es invertible se tiene que

$$(g - \gamma)' = \left(\sum_{k \in S} d_k f(h_k \gamma) h'_k \mathbf{x}_k \right)^{-1} (T_{\mathbf{x}} - \sum_{k \in S} d_k f(h_k \gamma) \mathbf{x}_k) + O_p(1/n).$$

Entonces la diferencia entre el estimador de calibración y el total viene dado por:

$$T_{y_{cal}} - T_{yU} = \sum_{k \in S} d_k f(h_k g) y_k - \sum_{k \in U} y_k = \sum_{k \in S} d_k f(h_k g) e_k - \sum_{k \in U} e_k,$$

siendo $e_k = y_k - \mathbf{x}_k(\sum_U f'(h_j\gamma)p_j h_j' \mathbf{x}_j)^{-1} \sum_U f'(h_j\gamma)p_j h_j' y_j$.

Ahora bien, como $p_j = \frac{1}{f(h_j\gamma)}$ entonces $\sum_{k \in s} f'(h_j\gamma)h_j' e_k = O_p(\frac{N}{n^{1/2}})$.

Por tanto

$$\begin{aligned} T_{y_{cal}} - T_{y_U} &= \sum_{k \in s} d_k f(h_k\gamma) e_k - \sum_{k \in U} e_k + \sum_{k \in s} d_k (f(h_k g) - f(h_k\gamma)) e_k = \\ &= \sum_{k \in s} d_k f(h_k\gamma) e_k - \sum_{k \in U} e_k + (g - \gamma)' \sum_{k \in s} d_k f'(h_k\gamma) h_k' e_k + O_p(\frac{N}{n}) \\ &= \sum_{k \in s} d_k f(h_k\gamma) e_k - \sum_{k \in U} e_k + O_p(\frac{N}{n}). \end{aligned}$$

Entonces el estimador de calibración es consistente bajo el diseño.

Respecto a su **ECM**, teniendo en cuenta que la probabilidad de que dos unidades k y j sean seleccionadas conjuntamente viene dada por $\tilde{\pi}_{kj} = \pi_{kj} p_k p_j$ y denotando por $\tilde{\pi}_k = \pi_k p_k$, obtenemos que

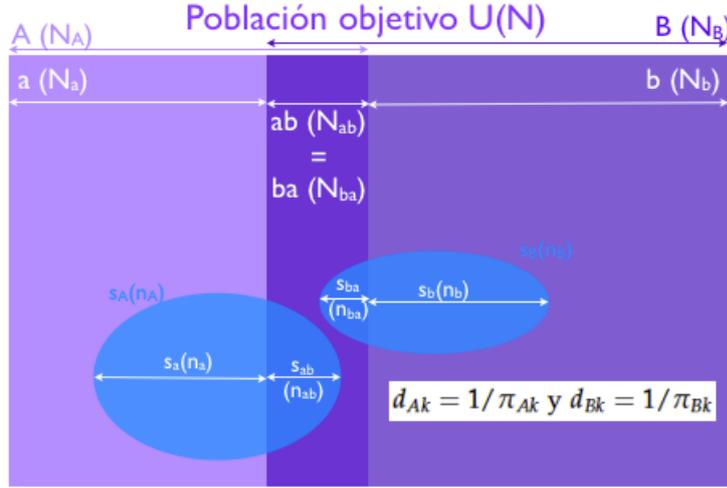
$$ECM(T_{y_{cal}}) = E[(T_{y_{cal}} - T_{y_U})^2] \simeq E[\sum_{k \in s} d_k e_k / p_k] = \sum_{k \in U} \sum_{l \in U} \Delta_{kl} \frac{e_k}{\tilde{\pi}_k} \frac{e_l}{\tilde{\pi}_l},$$

donde $\Delta_{kl} = \tilde{\pi}_{kl} - \tilde{\pi}_k \tilde{\pi}_l$.

6.3 VARIOS MARCOS IMPERFECTOS

Hay ocasiones en las que la elaboración de un único marco exhaustivo para la realización de encuestas es imposible, por lo que la solución es acudir a varios marcos disponibles que aumenten la cobertura poblacional. La resolución a este problema es lo que se conoce como encuestas con marcos múltiples cuyo desarrollo teórico fue realizado por **Hartley (1962)**. Así pues, presentó esta metodología como un procedimiento para reducir los costes de la recolección de datos al tiempo que se muestrea toda la población objeto de estudio, evitando así los sesgos derivados de la falta de cobertura. En su trabajo, Hartley demostró que las estimaciones calculadas combinando de forma adecuada la información recogida de diversos marcos de muestreo presentaban una eficiencia similar a las calculadas a partir de datos de un único marco muestral. Otra de las ventajas de las encuestas con marcos múltiples es que permiten utilizar un diseño muestral distinto en cada uno de los marcos. A partir de ese trabajo, han ido apareciendo otros muchos que

Figura 6.3: Contexto de trabajo en marcos duales.



han tratado este tema (ver [Lohr 2009](#) para una revisión de la estimación con marcos múltiples).

El procedimiento de estimación a partir de datos de marcos múltiples depende, principalmente, del número de marcos que se consideren y del grado de solapamiento entre ellos. A este respecto, la mayoría de los trabajos publicados hasta la fecha se centran en el caso de 2 marcos muestrales con solapamiento parcial entre ellos, que es lo que muestra la [Figura 6.3](#). A continuación se presenta la metodología de estimación en dos marcos y cómo se puede aplicar la técnica de calibración en este contexto.

6.3.1 Estimación en encuestas con dos marcos

Sea una población constituida por N unidades, $\mathcal{U} = \{1, \dots, k, \dots, N\}$. Sean \mathcal{A} y \mathcal{B} los dos marcos muestrales, ambos incompletos pero juntos cubren la población completa ([Figura 6.3](#)). La población de interés \mathcal{U} , puede dividirse en tres dominios excluyentes $a = \mathcal{A} \cap \mathcal{B}^c$, $b = \mathcal{A}^c \cap \mathcal{B}$ y $ab = \mathcal{A} \cap \mathcal{B}$. Dado que las unidades en el dominio de solapamiento ab pueden ser seleccionadas de una u otra muestra, se crea un dominio duplicado $ba = \mathcal{B} \cap \mathcal{A}$, idéntico a $ab = \mathcal{A} \cap \mathcal{B}$, para denotar las unidades duplicadas seleccionadas del marco B .

Sean $N, N_A, N_B, N_a, N_b, N_{ab}, N_{ba}$ los tamaños poblacionales de $\mathcal{U}, \mathcal{A}, \mathcal{B}, a, b, ab, ba$, respectivamente.

El total poblacional $T_y = \sum_{k=1}^N y_k$ de y , puede ser escrito como

$$T_y = T_{y_a} + \eta T_{y_{ab}} + (1 - \eta) T_{y_{ba}} + T_{y_b}, \quad (6.1)$$

donde $0 \leq \eta \leq 1$, $T_{ya} = \sum_{k \in a} y_k$, $T_{yab} = \sum_{k \in ab} y_k$, $T_{yba} = \sum_{k \in ba} y_k$ e $T_{yb} = \sum_{k \in b} y_k$.

Dos muestras aleatorias simples s_A and s_B son seleccionadas independientemente de A y B con tamaños n_A y n_B , respectivamente. Cada diseño induce probabilidades de primer orden π_{Ak} y π_{Bk} , y pesos muestrales $d_{Ak} = 1/\pi_{Ak}$ y $d_{Bk} = 1/\pi_{Bk}$, respectivamente, y que son definidos en función del dominio al que pertenezca la unidad k correspondiente. Las unidades en s_A se pueden expresar como $s_A = s_a \cup s_{ab}$, donde $s_a = s_A \cap a$ y $s_{ab} = s_A \cap (ab)$. Similarmente, $s_B = s_b \cup s_{ba}$, donde $s_b = s_B \cap b$ y $s_{ba} = s_B \cap (ba)$.

Un estimador del total viene dado por

$$\hat{T}_{yH}(\eta) = \hat{T}_{ya} + \eta \hat{T}_{yab} + (1 - \eta) \hat{T}_{yba} + \hat{T}_{yb}, \quad (6.2)$$

donde $\hat{T}_{ya} = \sum_{k \in s_a} d_{Ak} y_k$ es el estimador de Horvitz-Thompson del total del dominio a y similarmente para los otros dominios. [Hartley \(1962\)](#) propone elegir η de forma que minimice la varianza del estimador. Sea

$$d_k^\circ = \begin{cases} d_{Ak} & \text{si } k \in s_a \\ \eta d_{Ak} & \text{si } k \in s_{ab} \\ (1 - \eta) d_{Bk} & \text{si } k \in s_{ba} \\ d_{Bk} & \text{si } k \in s_b \end{cases}, \quad (6.3)$$

entonces $\hat{T}_{yH}(\eta) = \sum_{k \in s} d_k^\circ y_k$. En adelante eliminamos (η) para simplificar la notación.

Para η fijo, \hat{T}_{yH} es un estimador insesgado del total T_y . Puesto que los marcos A y B son muestreados independientemente, la varianza de \hat{T}_{yH} viene dada por

$$V(\hat{T}_{yH}) = V(\hat{T}_{ya} + \eta \hat{T}_{yab}) + V((1 - \eta) \hat{T}_{yba} + \hat{T}_{yb}). \quad (6.4)$$

ELECCIÓN PARA η . El problema de la elección más apropiada para η ha sido considerado por diversos autores. El valor de η que minimiza la varianza en [Ecuación 6.4](#) depende de varianzas y covarianzas que son desconocidas y que han de ser estimadas a partir de los datos, así como depende de la variable de interés. Esto origina la necesidad de recomputar los pesos para cada variable, lo cual es un inconveniente para las agencias de estadística y además produce inconsistencias en las estimaciones ([Lohr, 2009](#)).

El estimador desarrollado por [Fuller y Burmeister \(1972\)](#) incorpora la información de N_{ab} en la estimación. Sin embargo no es una combinación lineal de los valores de y salvo en [MAS. Skinner y J. Rao \(1996\)](#)

modifica el estimador propuesto por Fuller y Burmeister (1972) introduciendo un estimador de pseudo-verosimilitud máxima (PML) que no es óptimo pero puede escribirse como una combinación lineal de los valores observados y puede trabajar en diseños complejos.

J. Rao y C Wu (2010) extienden la metodología de pseudo-verosimilitud empírica (PEL) propuesta por C Wu y J. Rao (2006) para el caso de dos marcos, y define el estimador:

$$\bar{Y} = W_a \bar{Y}_a + W_{ab}(\eta) \bar{Y}_{ab} + W_{ba}(\eta) \bar{Y}_{ba} + W_b \bar{Y}_b, \quad (6.5)$$

donde $W_a = N_a/N$, $W_{ab}(\eta) = \eta N_{ab}/N$, $W_{ba}(\eta) = (1-\eta)N_{ab}/N$ y $W_b = N_b/N$, $\bar{Y}_{ab} = \bar{Y}_{ba}$, siendo $\eta \in (0, 1)$ una constante a especificar. Para ello se define la verosimilitud:

$$l_D(p_{ak}, p_{abk}, p_{bak}, p_{bk}) = n \left[W_a \sum_{k \in s_a} \tilde{d}_{ak} \log(p_{ak}) + \right. \\ \left. + W(\eta) \sum_{k \in s_{ab}} \tilde{d}_{abk} \log(p_{abk}) + \right. \quad (6.6)$$

$$\left. + W_{ba}(\eta) \sum_{k \in s_{ba}} \tilde{d}_{bak} \log(p_{bak}) + \right. \\ \left. + W_b \sum_{k \in s_b} \tilde{d}_{bk} \log(p_{bk}) \right], \quad (6.7)$$

donde $n = n_A + n_B$, $\tilde{d}_{ak} = d_{Ak} / \sum_{k \in s_a} d_{Ak}$, $\tilde{d}_{abk} = d_{Ak} / \sum_{k \in s_{ab}} d_{Ak}$, $\tilde{d}_{bk} = d_{Bk} / \sum_{k \in s_b} d_{Bk}$ y $\tilde{d}_{bak} = d_{Bk} / \sum_{k \in s_{ba}} d_{Bk}$. Los coeficientes se determinan maximizando la verosimilitud sujeta a las restricciones

$$\sum_{k \in s_a} p_{ak} = 1, \quad \sum_{k \in s_{ab}} p_{abk} = 1, \quad \sum_{k \in s_{ba}} p_{bak} = 1, \quad \sum_{k \in s_b} p_{bk} = 1,$$

junto a

$$\sum_{k \in s_{ab}} p_{abk} y_k = \sum_{k \in s_{ba}} p_{bak} y_k. \quad (6.8)$$

El estimador PEL es calculado de la forma

$$\hat{Y}_P = W_a \hat{Y}_a + W_{ab} \hat{Y}_{ab} + W_b \hat{Y}_b, \quad (6.9)$$

donde $\hat{Y}_a = \sum_{k \in s_a} \hat{p}_{ak} y_k$, $\hat{Y}_b = \sum_{k \in s_b} \hat{p}_{bk} y_k$ y $\hat{Y}_{ab} = \sum_{k \in s_{ab}} \hat{p}_{abk} y_k = \hat{Y}_{ba}$ dada la restricción Ecuación 6.8.

Por último, la metodología *single-frame* en encuestas de marcos duales combina las observaciones muestrales en una única muestra que ajusta los pesos del diseño para las unidades del dominio intersección. Para ello deben conocerse las probabilidades de inclusión de las unidades

del dominio ab para ambos diseños muestrales. Los pesos ajustados se definen de la forma siguiente:

$$d_k^* = \begin{cases} d_{Ak} & \text{si } k \in a \\ (1/d_{Ak} + 1/d_{Bk})^{-1} & \text{si } k \in ab \\ d_{Bk} & \text{si } k \in b \end{cases} . \quad (6.10)$$

El estimador propuesto por [Kalton y Anderson \(1986\)](#) es básicamente un estimador de Horvitz-Thompson:

$$\hat{T}_{yS} = \sum_{k \in S} d_k^* y_k. \quad (6.11)$$

Su varianza viene dada por $V(\hat{T}_{yS}) = V(\sum_{k \in S_A} d_k^* y_k) + V(\sum_{k \in S_B} d_k^* y_k)$. Si N_A y N_B son conocidos, este estimador puede ajustarse usando el método raking ratio ([Bankier, 1986](#); [Skinner, 1991](#)).

A continuación introducimos dos estimadores de calibración para marcos duales, dual-frame y single-frame.

6.3.2 Método de calibración dual-frame

Sea $\mathbf{x}_k = (x_{1k}, \dots, x_{pk})$ el vector de variables auxiliares con $T_x = \sum_{k=1}^N x_k$ conocido y donde, para cada dominio, obtendríamos $\mathbf{x}_{ak} = (x_{1ak}, \dots, x_{pak})$, $\mathbf{x}_{bk} = (x_{1bk}, \dots, x_{pbk})$ y $\mathbf{x}_{abk} = (x_{1abk}, \dots, x_{pabk})$. Usando la técnica de calibración vamos a modificar los pesos básicos de Hartley d_k° para obtener w_k° , para $k \in s$. [Ranalli y col. \(2015\)](#) definen el estimador de calibración dual-frame como

$$\hat{T}_{yCAL} = \sum_{k \in s} w_k^\circ y_k, \quad (6.12)$$

donde w_k° es tal que

$$\min \sum_{k \in s} q_k G(w_k^\circ, d_k^\circ) \quad \text{sujeto a} \quad \sum_{k \in s} w_k^\circ \mathbf{x}_k = T_x, \quad (6.13)$$

y $G(w, d)$ es una medida de distancia satisfaciendo las condiciones usuales de la calibración ([Deville y Särndal, 1992](#)). Notamos por $T_{xH} = \sum_{k \in s} d_k^\circ \mathbf{x}_k$ al estimador de Hartley de T_x . Entonces $w_k^\circ = d_k^\circ F(q_k \mathbf{x}_k \lambda)$, donde $F(u) = g^{-1}(u)$ y $g^{-1}(\cdot)$ denota la función inversa de $g(w, d) = \partial G(w, d) / \partial w$. El vector λ es calculado a partir de

$$\phi_s(\lambda) = \sum_{k \in s} d_k^\circ [F(\mathbf{x}_k \lambda) - 1] \mathbf{x}_k'$$

tal que $\phi_s(\lambda) = T_x - T_{xH}$.

Al usar diferentes medidas se pueden obtener diferentes estimadores de calibración. Bajo la distancia euclídea $q_k G(w_k^\circ, d_k^\circ) = (w_k^\circ - d_k^\circ)^2 / 2d_k^\circ$, se puede obtener la solución analítica $w_k^\circ = d_k^\circ(1 + \mathbf{x}_k \lambda)$, con $\lambda = (\sum d_k^\circ \mathbf{x}'_k \mathbf{x}_k)^{-1} (T_x - T_{xH})'$, y por tanto los pesos toman la forma

$$w_k^\circ = d_k^\circ \left[1 + (T_x - T_{xH}) \left(\sum_{k \in S} q_k d_k^\circ \mathbf{x}'_k \mathbf{x}_k \right)^{-1} q_k \mathbf{x}'_k \right]. \quad (6.14)$$

Así el estimador \hat{T}_{yCAL} puede escribirse como:

$$\hat{T}_{yCAL} = \sum_{k \in S} w_k^\circ y_k = \hat{T}_{yH} + (T_x - T_{xH}) \beta^\circ,$$

donde $\beta^\circ = (\sum_{k \in S} d_k^\circ q_k \mathbf{x}'_k \mathbf{x}_k)^{-1} (\sum_{k \in S} d_k^\circ q_k \mathbf{x}'_k y_k)$.

6.3.3 Método de calibración single-frame

Así pues, bajo el supuesto de que se pueda utilizar esta aproximación single-frame (Ecuación 6.11), el estimador de calibración viene dado por $\hat{T}_{yCAL}^S = \sum_{k \in S} w_k^* y_k$ donde los pesos w_k^* son tales que

$$\min \sum_{k \in S} q_k G(w_k^*, d_k^*) \quad \text{sujeto a} \quad \sum_{k \in S} w_k^* \mathbf{x}_k = T_x.$$

La diferencia con el estimador Ecuación 6.13 es el peso del diseño utilizado como base.

Si se utiliza la distancia euclídea,

$$\hat{T}_{yCAL}^S = \hat{T}_{yS} + (T_x - T_{xS}) \beta^*,$$

donde $\beta^* = (\sum_{k \in S} d_k^* q_k \mathbf{x}'_k \mathbf{x}_k)^{-1} \sum_{k \in S} d_k^* q_k \mathbf{x}'_k y_k$ y notamos por $T_{xS} = \sum_{k \in S} d_k^* \mathbf{x}_k$ al estimador single-frame de T_x .

Bajo ciertas condiciones dadas en Ranalli y col. (2015), \hat{T}_{yCAL}^S es un estimador consistente y asintóticamente equivalente al estimador de regresión en single-frame \hat{T}_{yGREG}^S . Entonces el estimador de calibración tiene su misma distribución asintótica por lo que su varianza se puede obtener a partir de la expresión $V(\hat{T}_{yGREG}^S) = V(\hat{t}_{eS}) = V(\sum_{k \in S} d_k^* e_k)$ que puede ser estimada usando $v(\hat{t}_{eS}) = v(\sum_{k \in S} d_k^* \hat{e}_k)$, donde $\hat{e}_k = y_k - \mathbf{x}_k \beta^*$.

Cuando N_a , N_b and N_{ab} son todos conocidos, si consideramos el vector auxiliar $\mathbf{x}_k = (\delta_k(a), \delta_k(ab) + \delta_k(ba), \delta_k(b))$, y, $T_x = (N_a, N_{ab}, N_b)$, la

solución no depende de la distancia elegida y los pesos de calibración adoptan la forma:

$$w_k^* = \begin{cases} d_{Ak} N_a / \hat{N}_a & \text{si } k \in s_a \\ (1/d_{Ak} + 1/d_{Bk})^{-1} N_{ab} / \hat{N}_{abS} & \text{si } k \in s_{ab} \cup s_{ba} \\ d_{Bk} N_b / \hat{N}_b & \text{si } k \in s_b \end{cases}, \quad (6.15)$$

donde $\hat{N}_{abS} = \sum_{k \in s_{ab} \cup s_{ba}} (1/d_{Ak} + 1/d_{Bk})^{-1}$.

7.1 INTRODUCCIÓN

A continuación se describen algunos de los programas informáticos específicos para aplicar la técnica de calibración de estimadores. La mayoría están desarrollados por las agencias estadísticas que los utilizan, aunque algunos, como los paquetes de R ([R Core Team, 2015](#)), son libres, por lo que no hace falta licencia, permisos, ni otros programas relacionados para utilizarlos.

7.2 PAQUETES ESTADÍSTICOS RELACIONADOS CON CALIBRACIÓN

Entre los programas estadísticos relacionados con calibración, se destacan los siguientes:

BASCULA

Desarrollado en Delphi por [Gottgens et al. \(1991\)](#) de [Statistics Netherlands](#), es uno de los paquetes más importantes utilizados para satisfacer la creciente necesidad de recopilar información estadística en encuestas por muestreo. BASCULA es parte del proyecto [Blaise](#), sistema integrado para el procesamiento de encuestas centrado en la entrada y edición de datos, entrevista (telefónica), captura electrónica de datos, gestión de encuestas, manipulación de datos y metadatos.

BASCULA está ya desarrollado en su versión 4.0, siendo una herramienta de fácil uso para el cálculo de pesos, basados en el diseño muestral y en la incorporación de la información auxiliar. Esta se recopila en un modelo de ponderación, que constituye la base para el procedimiento, estando los métodos basados en el estimador general de regresión (postestratificación, estimación ratio y ponderación lineal) y en la estimación raking ratio (ponderación multiplicativa). También proporciona ponderación consistente lineal entre personas y hogares, diagnósticos gráficos de los pesos y acotación de pesos.

BASCULA puede utilizar los pesos calculados para estimar totales poblacionales, medias y razones, así como varianzas basadas en la linealización de Taylor y/o en la replicación de muestras Balanced Repeated Replication ([BRR](#)), abarcando así los parámetros de población más comunes para las estadísticas oficiales. Una vez que los pesos finales se han calculado, se pueden aplicar directamente a un conjunto

arbitrario de variables objeto con el fin de obtener estimaciones de parámetros de población desconocida (más información en [Statistics Netherlands, Bascula 4](#)).

CALMAR

Fue el primer Software estadístico desarrollado para el método de calibración. Creado como una macro de Statistical Analysis System (SAS) por [Sautory \(1991\)](#) con el nombre de CALMAR (Calibración Marginal; CALage sur MARGes por sus siglas en francés), este Software produce pesos calibrados para las diferentes funciones de distancias descritas por [Deville and Särndal \(1992\)](#) ([Table 5.1](#)).

CALMAR es usado en la mayor parte de las encuestas sobre los modos de vida y presupuestos del Institute National de la Statistique et des Études Économiques (INSEE) de Francia, así como en la Encuesta de Población Activa del [Statistics South Africa](#), en los estudios del Censo Federal de Población (RFP) y la Encuesta sobre la Renta y el Consumo (ERC) realizadas por el [Statistique Suisse](#).

Su uso se ha extendido a otros organismos que producen estadísticas oficiales como el [INE](#), que lo utiliza en sus encuestas dirigidas a hogares, como la [EPA](#), la [ENSE](#) o la [ECV](#) ([C. Pérez and Portero, 2015](#)). Generalmente, las variables auxiliares que utiliza son los efectivos de población clasificados por grupos de edad y sexo, tratando de estandarizarlas para ofrecer resultados consistentes entre sus diferentes encuestas. En el caso de la [EPA](#), el [INE](#) ha optado por utilizar las proyecciones de población relativas a cada Comunidad Autónoma y referidas a cada trimestre.

Este programa permite utilizar una función de distancia entre cuatro opciones, asignándole al parámetro de entrada M los valores $M = 1, 2, 3$ y 4 . Si seleccionamos la función de distancia $M = 1$ da lugar al estimador GREG, en el caso de $M = 2$ da lugar al método raking para el caso de variables auxiliares cualitativas.

Es utilizado por numerosas oficinas de estadística públicas y centros privados de todo el mundo gracias a que, apoyándose en un desarrollo teórico riguroso, ha sido programado de forma que su utilización no es complicada, permitiendo además aplicar varias opciones de calibrado, tanto sobre variables auxiliares cualitativas como cuantitativas.

Se desarrolló una segunda versión, CALMAR2 ([Le Guennec and Sautory, 2004](#)), con nuevas incorporaciones que ayudan a mantener la consistencia en las estadísticas una vez calibrados los datos de las encuestas recogidas a diferentes niveles (por ejemplo, hogares e individuos). También corrige la falta total de respuesta propuesto por [Deville and Särndal \(1992\)](#), por lo que es posible volver a escribir las ecuaciones de calibración con dos conjuntos de variables: las variables de calibración actual y la falta de respuesta.

El programa CALMAR se puede descargar gratuitamente, junto con un manual de uso, de la página Web del [INSEE](#).

CLAN

Software desarrollado como una macro de [SAS](#) por [Statistics Sweden](#). Está concebido para el cálculo de estimaciones puntuales y errores estándar en encuestas a partir del método de linalización de Taylor. CLAN considera dos modelos diferentes de reponderación para el tratamiento de la no respuesta:

1. las respuestas independientes, con probabilidades de respuesta iguales dentro de los estratos, y
2. grupos de homogeneidad de respuesta en cada estrato (con la opción de dividirlos en dos o más grupos). Esta alternativa implica que CLAN puede ser utilizado también para un diseño muestral bifásico estratificado.

Más información en [Andersson and Nordberg \(2000\)](#).

G-CALIB

El paquete de Software g-CALIB fue realizado por [Vanderhoeft \(2001\)](#) en el [Statistics Belgium](#) y está disponible como módulo integrado en el paquete estadístico Statistical Package for the Social Sciences ([SPSS](#)).

Actualmente este paquete se encuentra en su versión 2.0 que se puede descargar de forma gratuita después de completar un formulario de inscripción en la Web del [Statistics Belgium](#).

g-CALIB es similar al CALMAR, aunque aparentemente algo más rápido. En [Chauvet et al. \(2005\)](#) puede encontrarse los resultados de un estudio comparativo de BASCULA, CALMAR y G-CALIB realizado por [Statistics Belgium](#).

GES (GENERALIZED ESTIMATION SYSTEM)

Desarrollado por [V Estevao et al. \(1995\)](#) para atender las necesidades de la estimación de muchas de las encuestas de [Statistics Canada](#). El enfoque de GES es la estimación por calibración mediante [GREG](#). El uso de información auxiliar en GES ha dado lugar a un moderno marco para la estimación, proporcionando cuatro funciones principales: cálculo de los pesos del diseño, el cálculo de g-pesos bajo un enfoque de calibración, el cálculo de las estimaciones de calibración y el cálculo de las estimaciones sintéticas, junto con la estimación de totales, promedios y proporciones, así como errores de muestreo. No es Software libre (más información en [Statistics Canada Quality Guidelines](#)).

GREGWT

Macro de SAS elaborada por [Australian Bureau of Statistics \(ABS\)](#). Realiza GREG y ponderación de los resultados de la muestra de una encuesta, así como estimación de la varianza mediante Bootstrap, Jackknife y BRR (más información en [Australian Bureau of Statistics](#)).

7.3 CALIBRACIÓN EN R

En esta sección se describen paquetes desarrollados en R ([R Core Team, 2015](#)) para la calibración de estimadores. R es un lenguaje y entorno de programación para análisis estadístico y gráfico. Se trata de un proyecto de Software libre, resultado de la implementación del lenguaje S. R y S-Plus (versión comercial de S) son, probablemente, los dos lenguajes más utilizados en investigación por la comunidad estadística, siendo además muy populares en el campo de la investigación biomédica, la bioinformática y las matemáticas financieras. A esto contribuye la posibilidad de cargar diferentes bibliotecas o paquetes con finalidades específicas de cálculo o gráfico, como los que se describen a continuación relacionados con las técnicas de calibración.

EVER

Este paquete, desarrollado por [Zardetto \(2012\)](#), permite estimar la varianza mediante replicación eficiente (Delete-A-Group Jackknife, [Phillip S Kott, 2008](#)). Incluye la función `kottcalibrate` ([Listing 7.1](#)) que añade al objeto construido con el diseño del estudio columnas con los pesos calibrados, una para cada peso replicado, y otra para los pesos iniciales.

Se puede ver la descripción de los argumentos en [Zardetto \(2012\)](#).

FRAMES2

Paquete desarrollado por [Antonio Arcos, Rueda, Ranalli, et al. \(2015\)](#) que permite realizar estimación puntual y por intervalos para encuestas con dos marcos muestrales. Implementa los clásicos estimadores de FB ([Fuller and Burmeister, 1972](#)), [Hartley \(1962\)](#), PEL ([J. Rao and C Wu, 2010](#)), PML ([Skinner and J. Rao, 1996](#)), SFRR ([Skinner, 1991](#)), así como dos estimadores de calibración: dual-frame y single-frame (ver [chapter 6](#) y [Antonio Arcos, Molina, et al. 2015](#)). El cálculo de las varianzas se puede hacer mediante las fórmulas asintóticas o mediante remuestreo (Jackknife). También permite dar los pesos de cada unidad de la muestra obtenida por los métodos de calibración anteriores.

Concretamente las funciones relacionadas con calibración son:

- `CalDF` ([Listing 7.2](#)): proporciona estimaciones para los totales y medias poblacionales utilizando el estimador de calibración DF

desde los datos muestrales obtenidos de un diseño muestral con dos marcos.

- `CalSF` ([Listing 7.3](#)): proporciona estimaciones para los totales y medias poblacionales, así como sus intervalos de confianza, utilizando el estimador de calibración SF desde los datos muestrales obtenidos de un diseño muestral con dos marcos.
- `JackCalDF` ([Listing 7.4](#)): calcula intervalos de confianza para el estimador de calibración dual-frame utilizando el procedimiento Jackknife.
- `JackCalSF` ([Listing 7.5](#)): produce estimaciones para la varianza del estimador de calibración SF mediante el procedimiento Jackknife.
- `WeightsCalDF` ([Listing 7.6](#)): calcula los g-pesos para el estimador de calibración dual-frame.
- `WeightsCalSF` ([Listing 7.7](#)): calcula los g-pesos para el estimador de calibración SF.

Se puede ver la descripción de los argumentos de las funciones anteriores en [Antonio Arcos, Rueda, Ranalli, et al. \(2015, págs. 5, 8, 26, 29, 55 y 57, respectivamente\)](#).

LAEKEN

Este paquete fue desarrollado por [Alfons et al. \(2015\)](#) y proporciona estimaciones de indicadores sobre pobreza y exclusión social, así como modelos de Pareto para distribuciones empíricas de ingresos. Tiene dos funciones relacionadas con calibración:

- `calibVars(x)`: función para construir una matriz de variables binarias a partir de x para calibrar los pesos muestrales según totales poblacionales marginales conocidos.
- `calibWeights` ([Listing 7.8](#)): calibra los pesos muestrales de acuerdo a totales poblacionales marginales conocidos. Basándose en los pesos muestrales iniciales, calcula los g-pesos mediante procedimientos de raking generalizado.

Se puede ver la descripción de los argumentos en [Alfons et al. \(2015, págs. 10 y 11\)](#)

Listing 7.1: Función kottcalibrate (paquete R: EVER)

```
kottcalibrate(deskott, df.population,
calmodel = if (inherits(df.population, "pop.totals"))
              attr(df.population, "calmodel"),
partition = if (inherits(df.population, "pop.totals"))
              attr(df.population, "partition") else FALSE,
calfun = c("linear", "raking", "logit"),
bounds = c(-Inf, Inf), aggregate.stage = NULL, maxit = 50,
epsilon = 1e-07, force.rep = FALSE)
```

Listing 7.2: Función CalDF (paquete R: Frames2)

```
CalDF(ysA, ysB, pi_A, pi_B, domains_A, domains_B, N_A = NULL, N_B =
      NULL, N_ab = NULL, xsAFrameA = NULL, xsBFrameA = NULL, xsAFrameB
      = NULL, xsBFrameB = NULL, xsT = NULL, XA = NULL, XB = NULL, X =
      NULL, met = "linear", conf_level = NULL)
```

Listing 7.3: Función CalSF (paquete R: Frames2)

```
CalSF(ysA, ysB, pi_A, pi_B, pik_ab_B, pik_ba_A, domains_A, domains_B
      , N_A = NULL, N_B = NULL, N_ab = NULL, xsAFrameA = NULL,
      xsBFrameA = NULL, xsAFrameB = NULL, xsBFrameB = NULL, xsT = NULL
      , XA = NULL, XB = NULL, X = NULL, met = "linear", conf_level =
      NULL)
```

Listing 7.4: Función JackCalDF (paquete R: Frames2)

```
JackCalDF(ysA, ysB, piA, piB, domainsA, domainsB, N_A = NULL, N_B =
          NULL, N_ab = NULL, xsAFrameA = NULL, xsBFrameA = NULL, xsAFrameB
          = NULL, xsBFrameB = NULL, xsT = NULL, XA = NULL, XB = NULL, X =
          NULL, met = "linear", conf_level, sdA = "srs", sdB = "srs",
          strA = NULL, strB = NULL, clusA = NULL, clusB = NULL, fcpA =
          FALSE, fcpB = FALSE)
```

Listing 7.5: Función JackCalSF (paquete R: Frames2)

```
JackCalSF(ysA, ysB, piA, piB, pik_ab_B, pik_ba_A, domainsA, domainsB
          , N_A = NULL, N_B = NULL, N_ab = NULL, xsAFrameA = NULL,
          xsBFrameA = NULL, xsAFrameB = NULL, xsBFrameB = NULL, xsT = NULL
          , XA = NULL, XB = NULL, X = NULL, met = "linear", conf_level,
          sdA = "srs", sdB = "srs", strA = NULL, strB = NULL, clusA = NULL
          , clusB = NULL, fcpA = FALSE, fcpB = FALSE)
```

REGENESEES

Este paquete R es el más reciente de entre los mostrados en esta tesis. Al igual que el paquete Ever, ReGenesees (R Evolved Generalized Software for Sampling Estimates and Errors in Surveys) fue realizado por [Zardetto \(2015\)](#) para el análisis basado en el diseño y asistido por el modelo de encuestas complejas. Es utilizado para la calibración, estimación y evaluación del error muestral en todas las macroencuestas llevadas a cabo por el *Istat* (Istituto nazionale di statistica, Italia). Se trata de la evolución y ampliación en R de la aplicación GENESEES realizada en SAS ([P. Falorsi y S. Falorsi, 1997](#)). De manera resumida:

1. La interacción del usuario con el nuevo sistema se produce en un muy alto nivel de abstracción. Así pues, los usuarios solo necesitan proporcionar los datos tal cual sean, y los metadatos simbólicos que describan el diseño muestral adoptado y el modelo de calibración, es decir, el modelo lineal asistido bajo un problema de calibración específico. Por lo tanto, el sistema transforma, automática y transparentemente, los datos muestrales en estructuras complejas de datos necesarios para resolver el problema de calibración y para calcular las estimaciones y errores.
2. Más allá de las distribuciones de frecuencia absolutas y relativas, ReGenesees permite calcular estimaciones y errores muestrales con respecto a las medias, razones, coeficientes de regresión múltiple, cuantiles, y, de manera más general, con respecto a cualquier estimador complejo, expresado como una función diferenciable de los estimadores de Horvitz-Thompson o de calibración. Así pues, el usuario solo necesita proporcionar al sistema la expresión simbólica del estimador como una función matemática. ReGenesees es de hecho capaz de linealizar esos estimadores complejos, de manera que la estimación de su varianza no suponga ningún coste al usuario.

En el trabajo de [Zardetto \(2015\)](#) se describe en profundidad las funcionalidades de este prometedor software.

SAMPLING

Este paquete fue desarrollado por [Tillé and Matei \(2015\)](#) para seleccionar y calibrar muestras. Así pues, permite extraer muestras complejas (muestreo con probabilidades desiguales, muestreo estratificado, muestreo por conglomerados, muestreo en dos fases), así como calcular los totales con calibración y con técnicas de regresión.

Utiliza varias funciones relacionadas con la calibración, entre las que destacamos:

- `calib` ([Listing 7.9](#)): Calcula los g-pesos del estimador de calibración. El g-peso debe estar entre los límites especificados para los métodos truncado y logit. La función devuelve el vector de g-pesos.
- `calibev` ([Listing 7.10](#)): Calcula el estimador de calibración y su estimación de la varianza utilizando la técnica de los residuos. La función devuelve dos valores: `cest` (el estimador de calibración) y `evar` (calcula su varianza estimada).
- `checkcalibration` ([Listing 7.11](#)): Comprueba la ecuación de calibración. En algunos casos, la calibración de los estimadores no existen, y los pesos calibrados (g-pesos) no permiten la calibración.
- `gencalib` ([Listing 7.12](#)): Calcula los g-pesos del estimador generalizado de calibración. Los g-pesos deberían estar entre los límites especificados para los métodos truncado y logit. Devuelve un vector con los pesos calibrados.

Se puede ver la descripción de los argumentos de las funciones anteriores en [Tillé and Matei \(2015, págs. 10, 12 y 8, respectivamente\)](#).

SURVEY

Este paquete creado por [Lumley \(2015\)](#) se utiliza en los análisis de muestras de encuestas complejas. Entre otras virtudes se caracteriza porque nos permite hacer estudios descriptivos, modelos lineales generalizados y estimaciones por máxima verosimilitud. Analiza diseños muestrales estratificados, por conglomerados, polietápicos y estudios con probabilidades desiguales. Además estima varianzas de estimadores por los métodos de linealización o por las técnicas de replicación ([BRR](#), Jackknife).

Relacionado con calibración, el paquete `survey` utiliza los métodos post-estratificación, raking y calibración lineal (o estimación [GREG](#)) que están relacionados con las formas de utilización de la información auxiliar disponible sobre el conjunto de la población. Estos métodos implican la adaptación de todos los pesos de muestreo a fin de que los totales de las variables auxiliares se reproduzcan exactamente.

El método de calibración está motivado por la estimación de una regresión total, pero puede ser calculado por reponderación simplemente, utilizando la función `calibrate`. Esta función, descrita en [Listing 7.13](#), implementa la calibración lineal, raking, y logit usando el algoritmo de Newton-Raphson descrito por [Deville, Särndal, and Sautory \(1993\)](#).

Listing 7.6: Función `WeightsCalDF` (paquete R: `Frames2`)

```
WeightsCalDF(ysA, ysB, pi_A, pi_B, domains_A, domains_B, N_A = NULL,
             N_B = NULL, N_ab = NULL, xsAframeA = NULL, xsBframeA = NULL,
             xsAframeB = NULL, xsBframeB = NULL, xsT = NULL, XA = NULL, XB =
             NULL, X = NULL, met = "linear")
```

Listing 7.7: Función `WeightsCalSF` (paquete R: `Frames2`)

```
WeightsCalSF(ysA, ysB, pi_A, pi_B, pik_ab_B, pik_ba_A, domains_A,
             domains_B, N_A = NULL, N_B = NULL, N_ab = NULL, xsAframeA = NULL,
             xsBframeA = NULL, xsAframeB = NULL, xsBframeB = NULL, xsT =
             NULL, XA = NULL, XB = NULL, X = NULL, met = "linear")
```

Listing 7.8: Función `calibWeights` (paquete R: `laeken`)

```
calibWeights(X, d, totals, q = NULL, method = c("raking", "linear",
        "logit"), bounds = c(0, 10), maxit = 500, tol = 1e-06, eps = .
        Machine$double.eps)
```

Listing 7.9: Función `calib` (paquete R: `sampling`)

```
calib(Xs,d, total,q= rep( 1, length(d)), method=c( "linear", "raking",
        "truncated", "logit"), bounds=c(low=0,upp=10), description=
        FALSE, max_iter=500)
```

Listing 7.10: Función `calibev` (paquete R: `sampling`)

```
calibev(Ys,Xs,total,pikl,d,g, q=rep(1,length(d)), with=FALSE, EPS=1e-6)
```

Listing 7.11: Función `checkcalibration` (paquete R: `sampling`)

```
checkcalibration(Xs, d, total, g, EPS=1e-6)
```

Listing 7.12: Función `gencalib` (paquete R: `sampling`)

```
gencalib(Xs,Zs,d,total,q=rep(1,length(d)),method=c("linear","raking",
        "truncated","logit"), bounds=c(low=0,upp=10),description=
        FALSE,max_iter=500,C=1)
```

Puede especificarse como una cadena de nombres con una de las tres funciones incorporadas o como objeto de la clase `calfun`, obtenido a partir de la función `make.calfun` que crea una métrica de calibración definida por el usuario.

Se puede ver la descripción de los argumentos en [Lumley \(2015, pág. 17\)](#)

TEACHINGSAMPLING

Este paquete fue desarrollado por [Gutiérrez \(2015\)](#) y permite al usuario seleccionar muestras probabilísticas y hacer inferencias desde poblaciones finitas basadas en distintos diseños muestrales. La función `Wk(x, tx, Pik, ck, b0)` calcula los pesos de calibración (distancia Chi-cuadrado) para la estimación del total poblacional para distintas variables de interés. Devuelve un vector con los pesos calibrados.

Se puede ver la descripción de los argumentos en ([Gutiérrez, 2015, pág. 83](#))

Una comparativa en cuanto a métodos disponibles y eficiencia de cálculo de la mayoría de estos paquetes se puede ver en [Antonio Arcos, Rueda, and Molina \(2013\)](#).

Listing 7.13: Función calibrate (paquete R: survey)

```
calibrate(design,...)

## S3 method for class survey.design2
calibrate(design, formula, population, aggregate.stage=NULL, stage
  =0, variance=NULL, bounds=c(-Inf,Inf), calfun=c("linear","
  raking","logit"), maxit=50, epsilon=1e-7, verbose=FALSE, force=
  FALSE, trim=NULL,...)

## S3 method for class svyrep.design
calibrate(design, formula, population, compress=NA, aggregate.index=
  NULL, variance=NULL, bounds=c(-Inf,Inf), calfun=c("linear","
  raking","logit"), maxit=50, epsilon=1e-7, verbose=FALSE, force=
  FALSE, trim=NULL, ...)

## S3 method for class twophase
calibrate(design, phase=2, formula,
  population, calfun=c("linear","raking","logit","rrz"),...)
grake(mm,ww,calfun,eta=rep(0,NCOL(mm)),bounds,population,epsilon
, verbose,maxit)
```


Part IV

APPLICATION OF CALIBRATION METHODS

Statistical agencies increasingly apply estimator calibration methodologies for their main surveys. This last part includes two chapters on the application of calibration in two real population-based surveys. Calibration is applied to estimate the prevalence of *infertility* in the first chapter ([chapter 8](#)) and to estimate the prevalence of *disabling chronic pain* in the second ([chapter 9](#)). Both chapters are structured according to the sections corresponding to an original scientific paper in the Health Sciences field. Accordingly, the first begins by introducing the theoretical framework of sterility and its prevalence and continues with the methodology and analysis of the *Infertility Survey* (Huelva, 2011) and the results obtained, finally analyzing the gain in accuracy after applying linear calibration. The second introduces the theoretical framework of chronic pain, its prevalence, and associated factors and continues with the methodology and analysis of the *Andalusian Health Survey* (2011) and the results obtained, analyzing the gain in accuracy for disabling chronic pain prevalence after applying linear calibration and the assisted model. Both chapters end with a discussion of the results and study limitations and strengths, and conclusions.

CALIBRATED PREVALENCE OF INFERTILITY

This chapter describes the implementation of the calibration techniques to estimate the prevalence of Infertility in 30- to 49-year-old women according to different approaches and from a cross-sectional population-based study (Huelva, Spain, 2011). Part of this work was published in *Human Reproduction* (A. Cabrera-León et al., 2015).

8.1 INTRODUCTION

Couples without reproductive problems who practice regular sex without birth control usually have an 85–90% likelihood of pregnancy within 1 year, reaching over 90% within 2 years (Australian Health Technology Advisory Committee 1998; Bedoya 1994; Guttmacher 1956; Maruani and Schwartz 1983). However, severe difficulties in achieving pregnancy are experienced by a substantive proportion of couples of childbearing age, estimated to be 10–15% by some authors (Navarro et al. 2006; Sexton et al. 2010) and 12–25% by others (Australian Health Technology Advisory Committee 1998; Bhattacharya et al. 2009; Boivin et al. 2007).

There is no clear consensus on the concept of infertility, which was defined as the inability to conceive after 1 year of regular intercourse and without using any type of contraception by the Sociedad Europea para la Reproducción Humana y Embriología (ESHRE) (The ESHRE Capri Workshop Group, 1994) and after 2 years, under the same conditions, by the World Health Organization (WHO) (Fuentes and Devoto, 1994). Taking account of the physical and emotional impact of infertility, it can be considered a chronic condition (Schmidt, Christensen, et al., 2005). Two distinct clinical situations are differentiated, designated primary and secondary infertility. Primary infertility refers to couples who do not achieve pregnancy after regular sexual intercourse without birth control for either 1 or 2 years, according to the definition, whereas secondary infertility refers to couples in this situation after a previous pregnancy without treatment (Larsen, 2005).

It has been recommended that research in this field includes self-reports to gauge individual perceptions of infertility and the potential demand for healthcare services (Larsen, 2005). Very few data are available on this issue in Spain, where it was reported that only one couple in every 1000 consults a doctor for this health problem (Peinado and

Bolúmar, 1997) and that not all couples with procreation difficulties seek medical advice (Jausoro, 2000).

Data from the 1999 fertility survey of the Spanish Statistical Office indicated that 2.4% of women aged between 15 and 49 years had undergone fertility treatment, most frequently in the 34–39 year age group, and that 51% of the treated women had failed to have children, whereas the remainder had given birth to at least one child (INE, 1999).

No data on the prevalence of infertility in southern Spain have been published by population studies, and information from other Spanish regions is .15 years old (Bolumar et al., 1996). This situation makes it difficult to estimate whether the planning of the relevant health resources in our country is appropriate to the demand, which is likely to be increasing, given the trend for couples to have a child at a later age and their greater awareness of the availability of effective therapeutic options (Jausoro 2000; Peinado and Bolúmar 1997). Information on the number of women without children after ending their fertile period is considered a biased estimation of the prevalence of infertility (Fuentes and Devoto, 1994) and is, therefore, not adequate to allow health providers to structure a care network that improves the reproductive health of the population. Consideration of infertility as a disease would support investigation into its causes and would make it more likely for women perceiving themselves in this situation to seek medical advice.

It is necessary for healthcare providers to establish the time to pregnancy before initiating infertility treatment in order to avoid over- or under-treatment. Gnoth et al. (2005) proposed the term subfertility (also known as subfecundity) to refer to a form of reduced fertility that prolongs the time until the desired conception, establishing three time periods (6, 12 and 48 months).

With this background, a population-based survey of women of child-bearing age was conducted in Southern Spain with the objective of estimating the prevalence of infertility from multiple approaches, including diagnosed and subjective infertility, primary and secondary 1-year infertility, and short-, medium- and long-term subfertility.

8.2 METHODS

8.2.1 Design

A descriptive cross-sectional population-based study was performed. Inclusion criteria were as follows: female sex, age between 30 and 49 years (of any sexual orientation), residence in the city of Huelva, Spain,

no difficulty in understanding Spanish and no severe cognitive impairment.

ETHICAL APPROVAL. This survey was reviewed and approved by the Accessibility and Continuity of Healthcare Board of the Andalusian Healthcare Public System (SAS, Spanish acronym) and the Andalusian School of Public Health (Escuela Andaluza de Salud Pública (EASP), Spanish acronym).

SAMPLE AND DATA COLLECTION. A sample size of 443 women was estimated, based on an assumed prevalence of infertility of 19% (Karmaus and Juul, 1999), a sampling error of +4.84 percentage points, a design effect of 1.8 (ratio of actual variance, under the sampling design used, to the variance of a simple random sample) and a confidence level of 95%. This sample size refers to the effective, valid, final sample. Therefore, a total of 2,371 households were selected to achieve the estimated sample size, and 3,543 contacts were made with candidate women (1.49 per household). The response rate was 61.05%.

A probabilistic, stratified and multistage sample design was used. The sample was distributed among each one of the 8 Huelva census districts (defined as strata) in proportion to the total study population (data from INE 2010; Table 8.1).

The primary sampling units were the census tracts of the city of Huelva (49.5% of the total census tracts were sampled; $N=107$; $n=53$). The secondary sampling units were households, which were selected through random routes (eight households per route and one route per census tract). When several women in the same household met with the inclusion criteria, the one who had most recently celebrated her birthday was selected. The mean duration of the interviews was 19 min (range, 7–52 min). In 77.4% of cases, only the interviewee and interviewer were in the house during the interview, ensuring a suitable environment for addressing the intimate and personal topics in the survey. The information was collected in March 2011. Interviews were carried out on any day of the week in the morning or evening, making it easier for all intended participants to be located.

After handing over the study presentation letter and obtaining the verbal informed consent of the women to participate in the study, information was gathered in face-to-face interviews by trained interviewers using a Computer-Assisted Personal Interview (CAPI) questionnaire. This questionnaire included a self-administered set of questions to collect the most personal information, ensuring the privacy and confidentiality of the data, enhancing the reliability of the results obtained and

Table 8.1: Distribution of the sample by age group, census district and census tracts.

Age group	Census district (number of census tracts sampled)	Proportion of the global population of Huelva city*	Interviews conducted
30-34	1 (8)	3.43%	13
	2 (6)	3.09%	14
	3 (5)	2.66%	13
	4 (6)	3.02%	11
	5 (6)	3.24%	12
	6 (5)	1.85%	9
	7 (12)	6.01%	30
	8 (7)	4.19%	17
35-39	1 (8)	3.49%	22
	2 (6)	3.47%	13
	3 (5)	2.53%	9
	4 (6)	2.81%	15
	5 (6)	2.69%	12
	6 (5)	2.01%	10
	7 (12)	5.45%	30
	8 (7)	3.22%	14
40-44	1 (8)	3.33%	11
	2 (6)	3.49%	14
	3 (5)	2.29%	11
	4 (6)	3.02%	9
	5 (6)	2.89%	12
	6 (5)	2.36%	11
	7 (12)	4.28%	22
	8 (7)	3.06%	15
45-49	1 (8)	3.43%	18
	2 (6)	2.49%	9
	3 (5)	2.11%	6
	4 (6)	2.78%	12
	5 (6)	2.61%	12
	6 (5)	1.97%	12
	7 (12)	3.85%	15
	8 (7)	2.88%	10
TOTAL**		100%	443

* INE 2010. ** All Huelva census districts and 49.5% (n=53) of the Huelva census tracts were sampled (1 route per census tract, 8 households per route and 1 woman per household)

maximizing the response rate. A pilot interview was conducted with 29 women.

OPERATIONAL DEFINITIONS. The final questionnaire was designed after analysis and evaluation of the pilot study results. It comprised interviewer administered sections on socio-demographic characteristics and reproductive history and a self-administered section on sexual habits and other intimate situations that could explain the failure to conceive.

Figure 8.1 and Table 8.2 define the eight different approaches of infertility considered in this study: diagnosed infertility, subjective infertility, 1-year infertility, primary infertility, secondary infertility and subfertility categorized according to the time taken to conceive (6, 12, or 24 months).

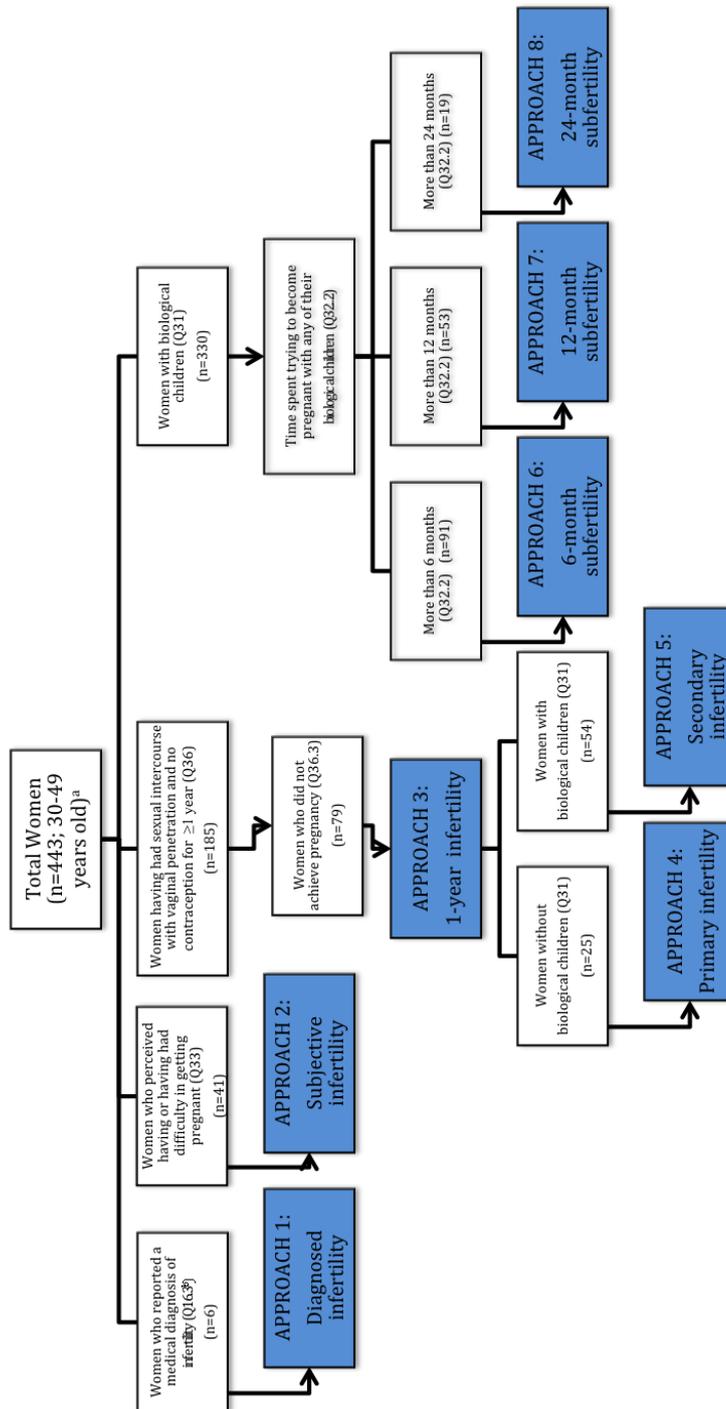
In order to reduce the high number of infertility categories to study the differences in socio-demographic features, several selection criteria were applied. In that sense, the Table 8.3 shows the selection criteria as well as the final selected infertility variables for each one; and the Table 8.4 shows the agreement coefficients for the selected variables in Table 8.3. Those criteria were established in pursuit of assuring a higher accuracy and providing different approaches to measure the prevalence of infertility. Thus, Subjective infertility, 1-year Infertility and 6-month Subfertility were finally selected.

8.2.2 Statistical analysis

Univariate descriptive analyses of the variables were performed, taking account of the study design and including the estimators for the means, proportions and totals, with 95% confidence interval (CI) and coefficient of variation (CV). Student t-tests and chi-square tests were used to study the differences in socio-demographic features among the three infertility categories finally selected.

The findings were extrapolated to the whole of Spain using the linear calibration approach (A Arcos et al., 2014; Särndal, 2007; Särndal and Lundström, 2005). Figure 8.2 shows the steps we took to select the auxiliary variables to be applied to calibration. So, the auxiliary data incorporated in the estimations, the calibration variables, were as follows: age, marital status, employment, educational level, country of birth and number of children (INE, 2015b). Because of confidentiality concerns, information on educational level and country of birth was not accessible for the city of Huelva; therefore, these characteristics could

Figure 8.1: Flowchart of the approaches employed in the study to estimate infertility prevalence.



a The boxes are not necessarily exclusive. In relation to maternity, approaches 1, 2 and 3 include women with or without biological children, approach 4 excludes women with biological children, while approaches 5 to 8 exclude women without biological children. Approaches 4 and 5 are mutually exclusive.

b Question codes; see questionnaire at <https://goo.gl/KsBn9o>.

Table 8.2: Approaches employed in the study to estimate Infertility prevalence.

<i>Situation in relation to maternity</i>	<i>Approach</i>	<i>Definition (question code; questionnaire in supplementary material)</i>
	1. Diagnosed infertility	Women with a medical diagnosis of infertility (question 16.3).
Women with or without biological children	2. Subjective infertility	Women who perceived having or having had difficulty in getting pregnant (question 33).
	3. Infertility 1 year	Women who failed to become pregnant after a year of vaginal sexual intercourse without contraception (questions 36 and 36.3).
Women without biological children (question 31)	4. Primary infertility	Women who did not achieve pregnancy despite trying it for over a year of sexual intercourse with vaginal penetration and no contraception (question 36.3).
	5. Secondary infertility	Women wanting to become pregnant again but failing despite trying it for over a year of sexual intercourse with vaginal penetration and no contraception (question 36.3).
Women with biological children (question 31)	6. Subfertility 6 months	Women having spent <u>more than 6 months</u> trying to become pregnant with any of their children.
Time taken to become pregnant with any of their biological children (question 32.2)	7. Subfertility 12 months	Women having spent <u>more than 12 months</u> trying to become pregnant with any of their children.
	8. Subfertility 24 months	Women having spent <u>more than 24 months</u> trying to become pregnant with any of their children.

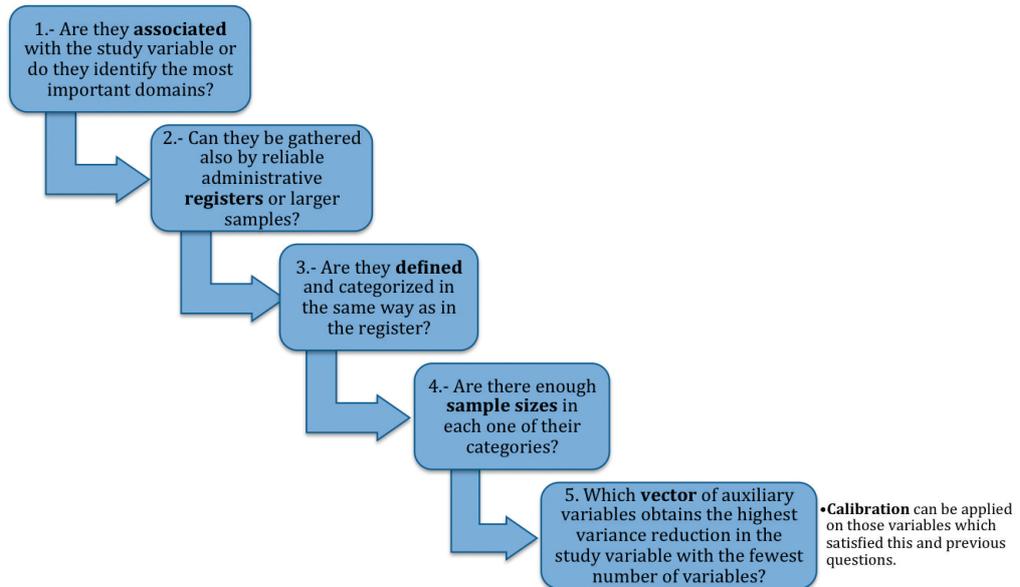
Table 8.3: Selection criteria of infertility variables to study the differences in socio-demographic features.

Criteria	Selected variables
1. Coefficient of variation under 20% (in order to assure an good accuracy on the correspondent estimation).	Subjective infertility; Infertility 1 year, Subfertility 6 months y 12 months
2. Lowest agreement with the rest of infertility variables (in order to provide a different point of view).	Subjective infertility; Infertility 1 year, Subfertility 6 months

Table 8.4: Agreement (Cohen's Kappa coefficient) for the prevalence estimators of infertility which satisfied the first selection criteria (see [Table 8.3](#)).

Variables	Subjective Infertility	Infertility 1 year	Subfertility 6 months	Subfertility 12 months
Subjective Infertility	-	0.051	0.288	0.317
Infertility 1 year	0.051	-	-0.154	-0.057
Subfertility 6 months	0.288	-0.154	-	0.669
Subfertility 12 months	0.317	-0.057	0.669	-

Figure 8.2: Questions to be satisfied by the calibration variables.



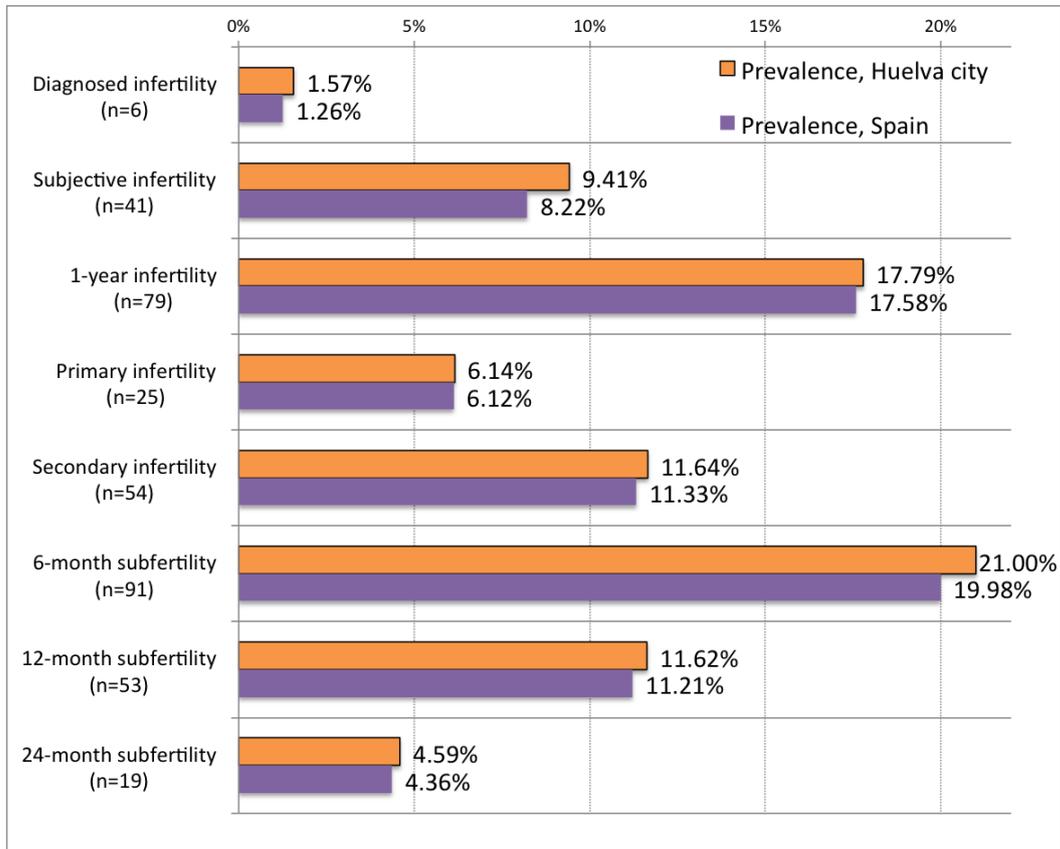
not be considered in the comparison between Huelva and Spain or for the estimator calibration.

The significance level was established at 0.05, and R software (R Core Team, 2015) was used for the data analyses. The ‘**sampling**’ (Listing 7.9) and ‘**samplingVarEst**’ (Lopez Escobar and Barrios Zamudio, 2015) R packages were used for the sample design and calibration weightings in all analyses.

8.3 RESULTS

PREVALENCE OF INFERTILITY. Figure 8.3 shows the estimated prevalence of infertility for Spain as a whole and Huelva city according to the different approaches used. For Spain, the prevalence among 30- to 49-year-old women with a medical diagnosis of infertility was 1.26% (95% CI: [1.26–1.27%]), and the difficulty in becoming pregnant at some time in their reproductive life (subjective infertility) was 8.22% (95% CI: [8.21–8.23%]). The prevalence of 1-year infertility (women who had not achieved pregnancy after 1 year of trying, including those who subsequently or never became pregnant) was 17.58% (95% CI: [17.57–17.59%]). The prevalence of primary infertility (failure to have a biological child) was 6.12% (95% CI: [6.12–6.12%]) and that of secondary infertility (failure to have another biological child) was 11.33% (95% CI: [11.32–11.37%]). The prevalence of 6-month subfertility (successful pregnancy not achieved within 6 months of starting to attempt conception among women with

Figure 8.3: Prevalence of infertility in 2011 for the study population (Huelva city, Spain) and following extrapolation to the entire Spanish population.



Estimations based on the sample design and on the entire study population (women aged between 30 and 49 years; n=443). The calibration variables for the Spanish prevalences were: age, marital status, employment, educational level, country of birth and number of children.

or without biological children) was 19.98% (95% CI: [19.97–20%]), whereas that of 12-month subfertility was 11.21% (95% CI: [11.2–11.22%]) and that of 24-month subfertility was 4.36% (95% CI: [4.35–4.37%]).

The population estimations obtained for infertility in Spain as a whole were as follows: diagnosed infertility in 95,182 Spanish women (aged between 30 and 49 years), subjective infertility in 845,027, 1-year infertility in 1,324,447, primary infertility in 461,284, secondary infertility in 853,565, 6-month subfertility in 1,505,887, 12-month subfertility in 845,027 and 24-month subfertility in 328,616.

SOCIO-DEMOGRAPHIC FEATURES. The mean age of the 443 women interviewed in Huelva city was 38.7 years (95% CI: [37.9–39.3 years]; range, 30–49 years); 94.7% were born in Spain, 39.3% had only primary schooling or less, 65.3% were married, 60.4% were in paid em-

ployment and 73% were mothers with a mean of 1.29 biological children (Table 8.5). There were no statistically significant differences in socio-demographic features among the women with subjective infertility, 1-year infertility or 6-month subfertility, except in the number of biological children (lower in the women with 1-year infertility and higher in those with 6-month subfertility) and in marital status (among the subpopulation with subjective infertility, fewer in the single and divorced categories and more in the married categories). Similar patterns of socio-demographic features were found among these three categories of women except for educational level, with a higher percentage of women with university studies in the subgroup with subjective infertility (42.9%) and a higher percentage of women with primary schooling or less in the subgroup with 6-month subfertility (45.8%). These differences suggested a trend but they were not statistically significant.

CHANGES AFTER CALIBRATING. As explained earlier, the prevalence and total estimations were calibrated for Spain alone, because of major socio-demographic differences between Huelva city (the sampling framework) and Spain as a whole. The largest differences were in employment (28% higher unemployment rate in Huelva than that in Spain) and number of children (24% fewer women with four or more children in Huelva). The calibration of prevalences reduced their estimated percentages and variances (Figure 8.4). The largest reduction in prevalence after calibration was for diagnosed infertility (-19.37%) and the smallest was for primary infertility (0.37%); the largest reduction in variance was for primary infertility (-90.46%) and the smallest was for 24-month subfertility (-3.7%). All calibrated prevalences had a CV of < 10%, guaranteeing adequate precision for the estimations except in the case of diagnosed infertility (CV of 17.7%).

8.4 DISCUSSION

This study investigated the prevalence of women with difficulties in getting pregnant and showed that this differed widely as a function of the approach to fertility measurement adopted.

Women under 30 years of age were not selected for this study because the mean age at first marriage in Spain was 33.6 years for men and 31.4 years for women, and the mean age of motherhood was 31.5 years overall and 30.1 years for the first child (INE, 2015b). It was therefore assumed that concern about a failure to become pregnant would not commence before the age of 30 years. In addition, a cohort study (Oakley et al., 2008) found a trend for women born later in the 1900s to seek a consultation regarding possible infertility at an older age. Thus,

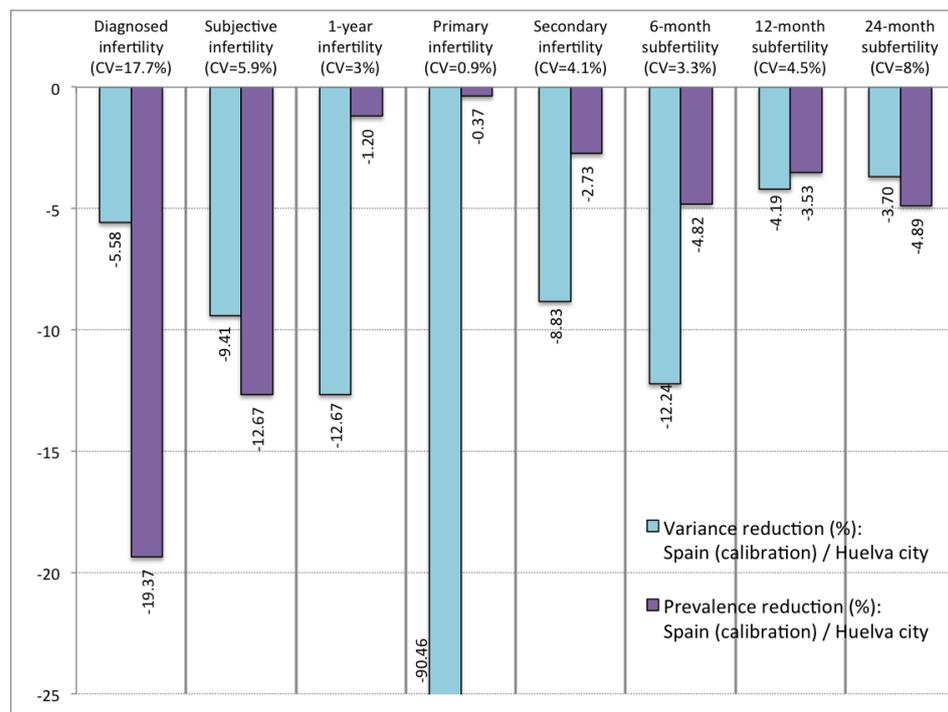
Table 8.5: Socio-demographic features of the study population (Huelva city, Spain, n=443 women) as a function of their infertility condition.

Description	Women of childbearing age (n=443)			Women with Subjective infertility (%)	Women with 1-year Infertility (%) ^b	Women with 6-month Subfertility (%) ^b
	%	95% CI (%)	CV ^a (%)			
Mother of biological children	73.0	67.7 – 77.7	3.4	76.3	65.5 ^c	100
Born in Spain	94.7	91.5 – 96.7	1.3	96.6	97.1	95
Educational level						
Primary schooling or less	39.3	33.3 – 45.7	7.9	34.5	32	45.8
Secondary schooling	28.0	23.0 – 33.6	9.5	22.5	38	24.8
University studies	32.7	27.5 – 38.3	8.3	42.9	30	29.4
Marital Status						
Single	21.1	17.1 – 25.7	10.1	2.7 ^c	17.8	4.6
Married / <i>De facto</i> partner / (<i>living</i> with partner)	65.3	60.0 – 70.3	3.9	94.3	74.4	80.8
Divorced / Separated (<i>not</i> living with partner)	11.5	9.1 – 14.4	11.5	3	7.3	11.6
Widow (not living with partner)	2.2	1.0 – 4.4	35.6	0	0.5	3
Employment Status						
Employed	60.4	53.8 – 66.7	53.0	47.7	58	59.8
Unemployed but previously employed	27.2	22.3 – 32.6	9.4	37.4	30.9	29.1
Seeking first job	2.0	0.7 – 5.9	55.1	0	1.7	1.3
Housework (not seeking employment)	8.9	6.4 – 12.4	16.5	14.9	7.9	8.6
Student (not seeking employment)	0.8	0.3 – 2.2	48.5	0	1.5	1.3
Disability / permanent disability	0.6	0.2 – 2.1	59.7	0	0	0
	Mean	CI (95%)	CV^a (%)	Subjective infertility (mean)	Infertility 1-year (mean)^b	6-month Subfertility (mean)^b
Number of biological children	1.29	1.17 – 1.41	4.6	1.23	1.1 ^c	1.84 ^c
Age (years)	38.7	37.9 – 39.3	0.9	38.4	37.8	40

a CV: Coefficient of variation;

b Subpopulation of women with the corresponding infertility condition;
 c p<0.001, Chi-square test and Student t-test; CI: confidence interval

Figure 8.4: Percentage reductions in variance and prevalence estimations after calibration



According to the largest reduction in variance estimators: age, marital status, employment, educational level, country of birth and number of children for Diagnosed infertility and Subjective infertility; age, marital status, employment, educational level and country of birth for Primary infertility and age, marital status, employment and educational level for the rest of the infertility categories. CV, Coefficient of variation for Spanish estimations.

for women born between 1945 and 1949, the mean age at consultation was 28.4 years for all who had consulted a physician and 29.1 years for those who had received treatment, compared with 30.8 and 30.5 years, respectively, for women born between 1960 and 1962. Likewise, in a retrospective study of a population of 1,043,513 UK inhabitants (Wilkes et al., 2009), the mean age of women consulting for infertility in a sample of couples failing to become pregnant after >1 year was 31 years. This suggests that women younger than 30 years are more likely to have contraceptive than proconceptive concerns. Women older than 49 years were not included because they may have a possible recall bias. This age group (30–49 years) also matches the one used by National Statistical Institutes, facilitating calculations and population comparisons. Furthermore, a similar age group (31–50 years) was studied in a comparable investigation in Scotland (Bhattacharya et al., 2009).

Special efforts were made to avoid sampling biases in this study, including: measures to minimize the effects of non-coverage, by selecting the study population within a sampling framework based on census districts and households and weighting the sample; measures to minimize nonresponse, by carrying out interviews 7 days per week between 10:00 and 21:00, using female interviewers trained in field work and the study methodology and furnished with the corresponding manuals, following up and supervising the survey administration, rewarding participation with a post-interview gift and replacing non-responders with women of the same age from the same district in a randomized manner (thus minimizing the differences between the effective and missing samples); measures to minimize information/observation/measurement biases, by providing adequate training to the interviewers (see above), using a CAPI system, performing a pilot field study, supervising the interviews with follow-up telephone calls, designing the questionnaire with filters and controls to facilitate automatic verification (by interviewer and interviewee) of its correct completion, incorporating self-administered sections for the more sensitive and intimate items and considering the sampling design in the data analyses.

Our response rate was acceptable in comparison with other studies (Thoma et al., 2013) developed by Official National Institutions (such as the National Statistic Institute in Spain), which encourage participation and have access to complete census registers, making the selection of participants easier and much more effective. Moreover, our survey may be less attractive to potential participants because of the specific and intimate information on infertility gathered. In fact, our response rate was higher than is usually reported for face-to-face interviews in population surveys focusing on infertility (Bhattacharya et al., 2009). It is likely that a lower response rate would have been obtained if we had

not introduced measures to minimize non-response, as detailed earlier. That response rate is different from that related to the missing values in each one of the infertility variables or categories, which were in all of them lower than 2%. So, we did not apply calibration to solve non-response problems because they were almost non-existent.

An increasing proportion of couples of childbearing age are unable to achieve a pregnancy despite regular unprotected sexual activity (Boivin, 2003). This situation produces anxiety and can have a negative effect on the relationship between partners (El-Messidi et al. 2004; Peterson 2006). Establishing the true magnitude of infertility as a public health problem is challenging, given that it is not categorized as a disability or chronic condition and may largely go unreported (Menken and Larsen, 1994). Although not universally acknowledged as such, the condition of infertility meets the requirements proposed by the WHO for the recognition of a disease in terms of its impact on physical, mental and social wellbeing. Unlike some other chronic diseases, infertility does not necessarily compromise life functions or produce physical limitations, and individuals can freely choose to undergo treatment or not, given that its progression is not life-threatening. Nevertheless, couples with fertility problems must face difficult consequences that can lead to the loss of dignity, relationship failure or even suicide (Ombelet et al., 2008). In fact, their emotional reactions to infertility have even been compared with those of cancer patients (Domar et al., 1993) and can worsen over time, being exacerbated by the diagnostic and treatment procedures involved. It should also be taken into account that the demand for care by affected couples increases with higher age and length of the period of infertility, whereas the effectiveness of therapy reduces with reproductive aging (de Graaff et al., 2011).

Furthermore, differences in the criteria used to define infertility hamper the comparison of results among epidemiological studies. Nevertheless, knowledge of the prevalence and risk factors of infertility in the community is necessary for health providers to assess the need for infertility treatment. According to Mahmoud Fathalla, at time of writing Chair of the WHO Advisory Committee on Health Research (Vayena et al., n.d.), the challenge of the new millennium is to achieve adequate access to services for the treatment of infertility, which requires knowledge not only of its incidence but also of the care demand that it generates. In this context, the time taken by couples to achieve a pregnancy appears to be a useful measure for estimating the prevalence of reproduction-related problems.

The present population-based study design provided an estimation of infertility according to different criteria, yielding data on medical diagnoses of infertility, the self-perception of the women, their sexual

relationships, the time required to conceive and the number of their biological children. This approach reveals the proportion of would-be mothers who manage to become pregnant in the short-, medium- and long-term, simplifying the comparison of results with the findings of other studies on reproductive/general health with varied time frames and criteria. The terms used vary among studies. For example, the category of 1-year infertility in the present study, i.e. the failure to become pregnant within 12 months (medium-term), is considered infertility per se or subfertility by different researchers.

Although we recorded the number of months that women reported trying to become pregnant with each of their biological children, we only asked whether they had practiced sexual intercourse with vaginal penetration and without contraceptive methods for .1 year. Hence, we could only calculate the time for which a couple were trying to become pregnant in the cases of women who finally became pregnant, and we do not know if it was longer than 1 year for the women failing to conceive. Future research should avoid this study weakness by gathering quantitative data on the months during which vaginal sexual intercourse was practiced, with no time limit.

The Spanish infertility prevalences provided by our study are comparable with those obtained in similar studies. Hence, the 1-year infertility prevalence (pregnancy not achieved within 1 year) in Spain (17.6%) is similar to findings in the USA (range: 12–18%) (Buck Louis et al. 2012; Thoma et al. 2013) and lower than that reported in France (24%) (Slama et al., 2012). The primary infertility prevalence (6.1%; proportion of women who attempted to have their first biological child but failed) is higher than the 4% observed by Bhattacharya et al. (2009). The 6-month subfertility prevalence (20%) is similar to the percentage reported in a Chinese population (22%) (Wang et al., 2003), whereas the 12-month subfertility prevalence (11.2%) is lower than previously described in Scotland (17.5%) (Bhattacharya et al., 2009), Finland (16.0%) (Terava et al., 2008), UK (17.3%) (Schmidt, Münster, et al., 1995), Denmark (15.7%) (Karmaus and Juul, 1999) or Spain (18.6% and 15.0%) (Karmaus and Juul 1999; Matorras Weinig 2011, , respectively). In contrast, the 24-month subfertility prevalence (4.3%) is higher than the percentage reported in Northwest Scotland (2.3%) (Bhattacharya et al., 2009). The only infertility prevalence with a high CV in the present study was for diagnosed infertility, which means that our estimations of its prevalence should be interpreted with caution.

A calibration of prevalence was needed to extrapolate the data to the whole of Spain because of the major differences in socio-demographic features between Huelva city and Spain (e.g. in employment rates and number of children). Without this calibration, all infertility preva-

lences would have been overestimated and less accurate. The calibration method offers a general framework for weighting that provides a simple and practical approach for incorporating auxiliary information into the estimation (Särndal, 2007). Calibration adjustments can be used to extrapolate the estimations of a survey, as in the present chapter, with the weightings providing sample ‘estimates’ for the totals of the auxiliary variables that match known population totals for these variables. In this way, the utilization of calibration estimators ensures that survey estimates are coherent with those already in the public domain, while simultaneously reducing sampling error and non-coverage bias (A Arcos et al. 2014; Särndal and Lundström 2005). Several national statistical agencies have developed software designed to compute weightings, usually calibrated in accordance with the auxiliary information available in administrative records and other accurate sources. An alternative technique to adjust weightings is the pseudo-empirical likelihood (PEL) method (J. Chen and R Sitter, 1999). Although the two approaches are asymptotically equivalent, PEL is not used in practice because of its complexity. In comparison with the PEL method, calibration is much easier to explain to users.

Similar patterns of socio-demographic features were found among the women with subjective infertility, 1-year infertility and 6-month subfertility (Table 8.5) except for the number of biological children, as expected, and educational level, as also reported by a recent comparative study of women in two infertility categories (Thoma et al., 2013). Further research is warranted into possible socio-demographic differences among the different infertility categories (Greil et al., 2010).

The 1-year infertility approach has also been proposed by other authors (termed ‘current duration approach at 12 months’) as a feasible alternative for estimating population-based infertility (Thoma et al., 2013). The time taken to become pregnant is a critical aspect of fertility success, and epidemiologic information on the proportions of women who fail to achieve pregnancy over the short-, medium- and long-term is critical for determining the timing of treatment initiation and the most appropriate therapy. Medical consultations for infertility have increased as couples more frequently attempt to start a family at older ages. From the age of 30 years onward, the pregnancy-related healthcare burden doubles for each 5-year delay in the women’s age (de Graaff et al., 2011). In fact, it has increased 4-fold over the past 20 years for women aged > 35 years, and the risk of not having the desired number of children is 10-fold higher among women aged > 30 years than among those aged 20–24 years (Joffe et al., 2009). In Spain, the pregnancy-related healthcare burden has augmented from 47.9% in 1995 to 63–70% in 2003 (López, 2014).

In conclusion, there are major differences in the prevalence of women with difficulties in getting pregnant according to the approach adopted. However, the different prevalences showed similar patterns of sociodemographic characteristics.

This chapter offers a set of categories to measure infertility, including subjective infertility, 1-year infertility and subfertility, which may be useful for cross-disciplinary comparisons of infertility in clinical and population-based studies. They may also assist health managers to tailor reproduction resources and services to the real needs of their populations and provide a more rapid and successful response to couples. Finally, the calibration adjustments applied to the infertility prevalences increased the validity of their generalization from a small sample and area (Huelva city) to a much larger one (Spain) and also improved their accuracy. This approach is being increasingly adopted in social sciences, especially in studies with small samples or with coverage and non-response problems. The methodology developed in our study may serve as a model for application in other epidemiological studies.

CALIBRATED PREVALENCE OF DISABLING CHRONIC PAIN

This chapter describes the implementation of the calibration techniques to estimate the prevalence of Disabling Chronic Pain in the general adult population from a cross-sectional face to face population-based study (Andalusia, 2011). Part of this work is being reviewed in Pain Medicine (Andrés Cabrera-León, Cantero-Braojos, et al., 2016) and in Pain (Andrés Cabrera-León, Rueda, et al., 2016).

9.1 INTRODUCTION

Estimations of the prevalence of Chronic Pain (CP) have varied widely among studies, mainly due to differences in definition, study population, design, sample size, and/or data-gathering procedure (Johannes et al., 2010; Kennedy et al., 2014; Langley et al., 2011; Toth et al., 2009). Thus, recent calculations of CP prevalence among adults (≥ 18 yrs) have ranged between 12 and 42% worldwide, between 12 and 30% in Europe (Breivik et al., 2006), and between 19 and 30.7% in the USA (Johannes et al., 2010; Kennedy et al., 2014). It was reported to be 35% in Canada (Toth et al., 2009), 18.5% in Australia (Blyth, March, AJ Brnabic, et al., 2006), 17.5% in Japan (Sakakibara et al., 2013), 35% in Hong Kong (Wong and Fielding, 2011), and 42% in Sao Paulo (Costa Cabral et al., 2014). CP prevalence has been found to range between 12 and 17.25% in Spain (Breivik et al., 2006; Dueñas et al., 2014; Langley et al., 2011) and between 15 and 18.3% in the Southern Spanish region of Andalusia (Herrera-Silva et al., 2012; Langley et al., 2011).

The mean age of patients with CP is around 50 yrs, and its prevalence is higher in females and individuals with lower socioeconomic level (Blyth, March, AJ Brnabic, et al., 2006; Breivik et al., 2006; Costa Cabral et al., 2014; Johannes et al., 2010; Langley et al., 2011; Toth et al., 2009) or living in an old people's home [1,9]. The impact of CP is greater when it limits Activities of Daily Living (ADL) (Azevedo et al., 2012; Blyth, March, AJ Brnabic, et al., 2006; Breivik et al., 2006; Langley et al., 2011; Raftery et al., 2011; Reid et al., 2015). The WHO includes disability-related ADL limitations within the "International Classification of Functioning, Disability, and Health" (ICF Model) (WHO, 2006). This biopsychosocial model considers disability as a state of impaired functioning associated with disease, disorder, lesion, or other health

condition that is experienced as a deficiency, a limitation on activity, or a restriction on participation in any area of life. There have been numerous studies on disability in different diseases, but few on its relationship with CP. The frequency of ADL-limiting CP or Disabling Chronic Pain (DCP) is higher in females and in individuals with lower socioeconomic level, health-related unemployment, disability/incapacity (Azevedo et al., 2012; Blyth, March, AJ Brnabic, et al., 2006), elevated depression indicators (Azevedo et al., 2012; Raftery et al., 2011; Reid et al., 2015), and/or higher number of visits to their physician (Blyth, March, AJM Brnabic, et al., 2014).

Most population health surveys on CP have considered it as a symptom of different chronic diseases, while others have considered CP as an independent entity and have associated it with various comorbidities (Dominick et al., 2014; Eriksen et al., 2003). These studies were based on allostatic load models (McEwen, 1998) and found that the capacity of individuals to adapt to stress factors can be impaired by the presence of CP and two or more comorbidities, increasing health risks.

One of the most recent techniques to make the estimation of prevalence more valid and accurate is the calibration (Särndal, 2007). This statistical method offers a general framework for weighting that provides a simple and practical approach for incorporating auxiliary information into the estimation. Calibration adjustments can be used to extrapolate the estimations of a survey, as in the present chapter, with the weightings providing sample 'estimates' for the totals of the auxiliary variables that match known population totals for these variables (A. Cabrera-León et al., 2015). In this way, the utilization of calibration estimators ensures that survey estimates are coherent with those already in the public domain, while simultaneously reducing sampling error and non-coverage or nonresponse bias (Lundström and Särndal, 2001). Several national statistical agencies have developed software designed to compute weightings, usually calibrated in accordance with the auxiliary information available in administrative records and other accurate sources (A Arcos et al., 2014).

The objectives of the initial study were to calculate the prevalence of DCP among adults in Andalusia (southern Spain) as a function of key sociodemographic characteristics, to determine its multimorbidity and relationship with other chronic diseases, and to identify associated factors, based on the findings of a wide cross-sectional population-based survey on health and healthcare services.

A new objective was derived from those objectives: to offer a more representative and accurate estimation of the prevalence of DCP in Andalusia by applying different calibration technique approaches.

9.2 METHODS

9.2.1 *Design*

The information source for this study was the 4th edition of the EAS (2011), a cross-sectional population-based survey (Sánchez Cruz et al., 2013) designed to evaluate the health of non-institutionalized adults (≥ 16 yrs) in Andalusia (southern Spain) and their utilization of health services.

A multi-stage stratified sample design was adopted. Sampling units were (stages 1 to 4): municipalities, census tracts, households, and individuals. Strata were based on the province (8 strata) and size of municipality (5 strata: municipalities up to 2,000 inhabitants, from 2,001 to 10,000, from 10,001 to 50,000, and above 50,000 inhabitants, and provincial capitals). Municipalities and census tracts were selected in proportion to the population size, while households were selected with equal probability by systematic sampling. Interviewees were selected by applying quotas for each province and quotas for sex-age and size of municipality within each province. A virtually constant assignment was performed per census tract (7-10 adults), and one adult per household was selected for interview.

ETHICAL APPROVAL. The EAS was supervised and approved by the review board of the General Secretariat of Quality and Public Health in the Health Ministry of the Andalusian Regional Government.

SAMPLE AND DATA COLLECTION. Possible seasonal effects were minimized by establishing a third stratification based on the time in which the interviews were conducted, considering four three-month periods; 43.1% of interviewees were followed up (in person or by telephone), and all questionnaires were reviewed for correct completion and internal consistency; 112 municipalities and 696 census tracts were selected, and 6,507 valid personal face-to-face interviews were conducted at home ($p=q=0.5$; confidence level = 95%; precision=0.0149; design effect =1.525).

VARIABLES. Table 9.1 shows the study variables.

9.2.2 *Statistical analysis*

DESCRIPTIVE ANALYSIS. For dependent variables and their crossing with independent variables (tab:Studyvariables), the tables report data on percentages, means, 95% confidence intervals, coefficients of varia-

Table 9.1: Study variables

Variable Role	Definition and observations	Code (a)
Dependent	Chronic Pain (CP): individuals who declared that a doctor or a nurse had told them that they suffered from one or more of the following Chronic Diseases in the survey that included the word 'pain': 'migraine/headache/chronic cephalalgia /frequent headache', 'angina/chest pain', 'back pain, neck pain, shoulder pain, waist pain, cervical/low back pain or menstrual pain'.	P.14
	Disabling Chronic Pain (DCP): Population with CP who declared being limited in their activity by any of the above-reported chronic pains. The non-disabling CP population was also defined with this variable	P.14b
	Pain interference with regular work: sub-question of the scale SF-12 (Vilagut et al., 2008). This variable cannot be used to estimate DCP because it is not possible to discriminate between patients with pain for more than 3 months and those whose pain appeared in the past week.	P.11
Independent (b)	Sex and age (age groups: 16-44 years; 45-64; + 65)	P.2
	Number of chronic diseases: cancer, diabetes, hypertension, high cholesterol, colitis and chronic intestinal disease/ inflammatory bowel disease; stomach ulcer; chronic constipation; chronic lung disease; asthma; cardiac disorder; heart attack; fibromyalgia; chronic skin problems; chronic allergy; anemia; poor circulation; varicose leg veins; hemorrhoids; stroke; depression or anxiety; other mental problems; hearing loss; cataract; arthritis or rheumatism; osteoporosis; cirrhosis; kidney disease; urinary incontinence; infertility; prostate disorder (men); thyroid diseases. Chronic diseases in the CP dependent variable were excluded.	P.14
	Physical and mental health-related quality of life (Vilagut et al., 2008).	P.7-P.13
	Question on self-rated health status in the last 12 months (Verbrugge, 1997).	P.40
	Functional social support: total score and as confidant and affective dimensions (Ayala et al., 2012).	P.57
	Demographic and economic characteristics: marital status, cohabitation, living alone at home, social class (Chilet-Rosell et al., 2012), educational level, employment status, economic difficulty to make ends meet, total revenues.	P.59/59a, 1, 63b, 60/61, 71/72
Independent (c)	Sleep and rest during sleeping hours	P.33/33a
	Limitation, disability or physical, sensory or mental handicap for more than 6 months	P.2aa
	Healthy eating habits as (Dapcich et al., 2013): ≥ 1.5 liters of water/day; milk, fruit, vegetables, fish, ≥ 3 times/week; bread and cereal, ≥ 1 times/week; legume, pasta, rice, potatoes, ≥ 3 times/week (without being daily) or < 1 /week; meat, ≥ 2 times/week (without being daily); sausage, 1-2 times/week or never/almost never; eggs: 1-2 times/week. Sweets: < 1 /week or never/almost never	P.35a/36
	Suspected alcoholism (Ewing, 1984), frequency of consumption of alcoholic beverages, tobacco consumption.	P.32/29/23
	Body Mass Index (BMI) as continuous variable and categorized as: low weight ($BMI < 18.5kg/m^2$); normal weight ($18.5kg/m^2 \leq BMI < 25kg/m^2$); overweight ($25kg/m^2 \leq BMI < 30kg/m^2$); obesity ($BMI \geq 30kg/m^2$)(WHO, 2015b). Both size and weight were measured objectively.	P.46/47
	Physical activity in the workplace (P.34) and physical exercise in free time	P.34a
	Environmental quality of the area of residence from responses to general self-assessment questions and items on noise, smell, air pollution, industry, green areas, delinquency/insecurity and heavy traffic. The sum of the scores for these items was calculated and then categorized into tertiles (q33.34=18; q66.66=19). Factorial analysis was also performed using these variables.	P.5b/5
	Physical work conditions (working population), The sum of the scores for the 7 items (Liker scale responses, 1 to 4) was calculated and then categorized into tertiles (q33.34=20; q66.6=24).	P.66
	Psychosocial level occupational exposure (Nübling et al., 2014) (working population). Two components: 1) psychological demands; 2) active work and development possibilities, such as influence, skill and time control. For both components, the sum of the scores for the corresponding items (Liker 1 to 5) and then categorized into 3 tertiles (q33=10 and q66=15, component 1; q33=26 and q66=34, component 2).	1) P.67.1-6 y 2) P.67.7-16
	Deprivation Index (municipality of residence of the respondent) constructed from the Census 2011 (INE, 2015b), in which the theoretical scores are distributed with a mean of 0 and standard deviation of 1, with a higher score indicating a more deprived socioeconomic situation (Sánchez-Cantalejo et al., 2008).	
Rurality Index (municipality of residence of the respondent) constructed from the Census 2011 with a theoretical score range from -3.50 to 3.78; a higher score indicates a greater degree of rurality(Prieto Lara and R, 2010).		

(a) Item code in the questionnaire of the Andalusian Health Survey (see it at <https://goo.gl/PC6umv>). (b) Variables used to determine the prevalence of the dependent variables in each stratum. (c) Used to analyze multiple effects in multivariate models.

tion (CVs), corrected typified residuals, and p-values obtained in the statistical tests (Pearson's chi-square test corrected with second-order Rao-Scott and Mann-Whitney U tests).

Factorial analysis was performed on environmental quality items for subsequent analysis of their association with CP, obtaining two main factors: bad odors and atmospheric pollution factors; and safety, noise, and green spaces.

FACTORS ASSOCIATED WITH DISABLING CHRONIC DISEASE. Multivariate multinomial logistic regression models were used to analyze the association of factors with DCP, non-disabling CP, and absence of CP. A model was initially adjusted using a backwards stepwise procedure, with sociodemographic variables as control variables along with remaining secondary variables. Those furthest from significance (at 5%) were successively and manually excluded, verifying at each step that the exclusion did not change the value of the other parameters by >30% of their previous value; variables were re-entered in the model as confounding variables if a change > 30% was observed (Miettinen and Cook, 1981). The effects of age and gender interactions with the remaining independent variables were also verified in the data modeling process, only considering those that were statistically significant in the final model. Model assumptions were verified using residuals, model convergence, continuous variable linearity, variations in estimation standard error, and Nagelkerke R-square values (Menard, 2000).

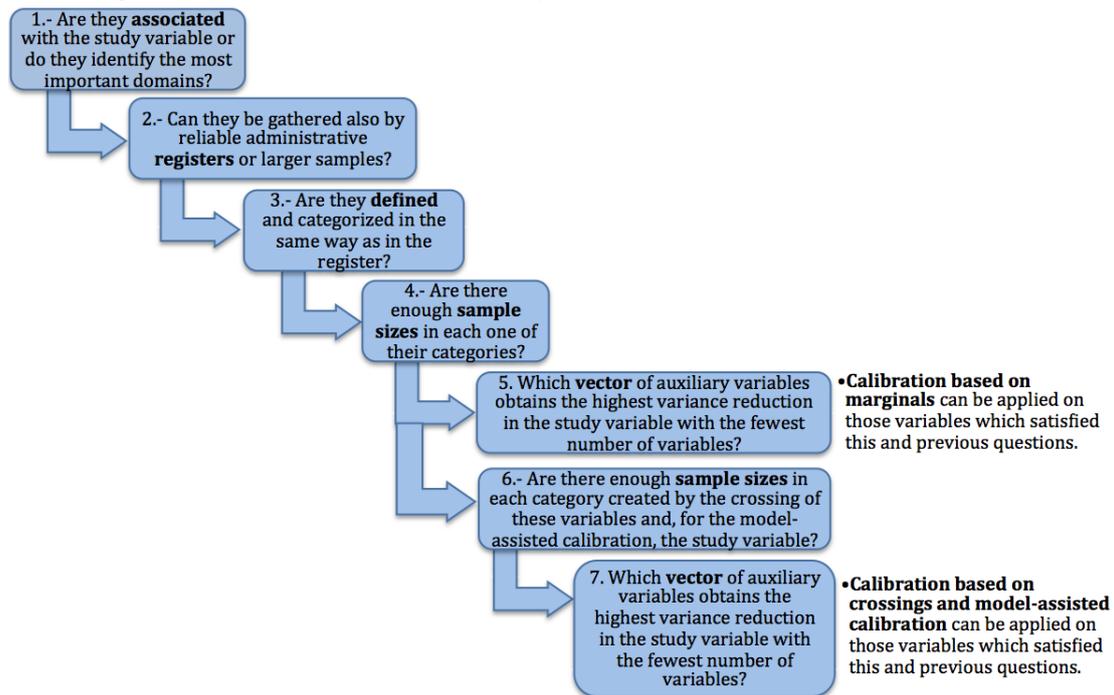
CALIBRATION. The following approaches were considered to estimate the prevalence of DCP according to sampling weights:

1. *Horvitz-Thompson* estimation (original sampling weights) (Sarjinder Singh, 2003),
2. linear calibration based on marginals (number of individuals in the population for each category of the calibration variables),
3. linear calibration based on crossings (number of individuals in the population for each category created by crossing the calibration variables), and
4. model-assisted calibration (Changbao Wu and Randy R Sitter, 2001).

Estimations on prevalence, population, and variance were calculated to compare the changes among these calibrations approaches.

The (first) condition to be met by *auxiliary calibration variables* was an association with the DCP study variable (Figure 9.1). Based on the final multivariate multinomial logistic model (from now on, general model),

Figure 9.1: Questions to be satisfied by the calibration variables.



a total of 11 variables with statistical significance <0.1 were identified as candidate calibration variables (see Figure 9.3). Among these variables, only sex and age could be obtained from the population records (INE, 2011), the (second) condition for calibration; hence, these would be the initial auxiliary variables to calibrate prevalence. The EAS (the study survey) is the most reliable source of information in Andalusia to estimate the remaining variables entered in the final model. Population variables for tobacco, hours of sleep, physical activity, affective social support, and physical limitation (Table 9.1), could also be obtained from the estimations provided by the ENSE but only for Spain as a whole and not for Andalusia, the geographic setting of the study; therefore, they were not considered for calibration. In addition, the variables physical and mental quality of life, environmental conditions, or number of chronic diseases were also obtained from the ENSE, although not with the same definition or categorization, the (third) condition required for calibration; hence, they also could not be considered as calibration variables.

Besides the global model, as reported in the Statistical Analysis paragraph (subsection 9.2.2), other multinomial logistic models were created, controlling for three variables (from now on, specific models): sex, age group (16-44, 45-64, and ≥ 65 yrs), and other variables not considered in the general model. Based on these models, we obtained

two variables with statistical significance < 0.1 that also had population record data and the same definitions: educational level ($p=0.056$) and employment status ($p<0.001$).

Other variables with data available from population records (INE, 2011) showed statistical significance >0.1 and were therefore not considered for calibration: caring for an under-15-yr-old, caring for a person with health problems, being responsible for most domestic chores, marital status, children, and the size, heating, and type of home.

The (fourth) calibration condition was to check whether there were at least two observations for every category of the calibration variable (for calibration approach 2) or for each category created by crossing the calibration variables with the study variable (for approaches 3 and 4). In our study, crossing categories were DCP, sex, age groups, and educational level (36 categories) and sex, age, and employment status (70 categories). Hence, only sex, age groups, and educational level had a sufficient number of observations and were considered as the final calibration variables.

Finally, the model-assisted calibration was implemented with a binary logistic multivariate model in which DCP was dependent variable (two categories: DCP vs. non-Disabling Chronic Pain (nDCP) or non CP) and calibration variables were independent variables. Probability-predicted values from this model were obtained for utilization in the model-assisted calibration.

Simple and stacked bar graphs and Odds Ratio (OR) synthesis graphs were created.

The significance level was established at 0.05, and SPSS and R Software (R Core Team, 2015) were used for the multivariate logistic models and calibration, respectively. The 'sampling' (ver Listing 7.9) and 'samplingVarEst' (Lopez Escobar and Barrios Zamudio, 2015) R packages were used for the sample design and calibration weightings in all analyses.

9.3 RESULTS

SOCIO-DEMOGRAPHIC FEATURES. Table 9.2 and Table 9.3 list the main sociodemographic, economic, and life habit characteristics of the study population and the number (and percentage) missing for each variable.

As shown in Table 9.4, CP was reported by 17.2% of the study population (1,187,616 individuals (95% Confidence Interval (CI)=[1,124,707-1,250,525])). The prevalence of CP was more than two-fold higher in females versus males (23.9% vs. 10.3%, $p<0.001$). The mean age of individuals with CP was 54.12 yrs (95%CI=[53.01-55.2]), more than 10 yrs

Table 9.2: Characteristics of the study population (Andalusia, 2011).

Variables (individuals with missing data; %)	Category (population aged ≥16 years; total n=6,507)	Percentage (n)	95% Confidence Interval	
			Minimum	Maximum
Sex (0)	Male	49.3% (3209)	48.1%	50.5%
	Female	50.7% (3298)	49.5%	51.9%
Age (0)	Aged 16-24 years	13.3% (867)	12.5%	14.2%
	25-44	39.3% (2556)	38.1%	40.5%
	45-54	16.7% (1087)	15.8%	17.6%
	55-64	12.3% (799)	11.5%	13.1%
	65-74	9.6% (626)	8.9%	10.4%
	+75	8.8% (572)	8.1%	9.5%
Marital Status (8, 0.1%)	Married	59.6% (3805)	58.4%	60.8%
	Single	29.6% (1882)	28.5%	30.7%
	Separated	1.9% (123)	1.6%	2.2%
	Divorced	2.8% (178)	2.4%	3.2%
	Widowed	8% (511)	7.4%	8.7%
	Cohabiting (66, 2.6%)	64.9% (4180)	63.7%	66%
Social Class [27] (1359, 20.9%)	I. Manager with 10<salaried staff	5.3% (275)	4.8%	6%
	II. Manager with 10>salaried staff	7.5% (384)	6.8%	8.2%
	IIIa. Administrative staff	12.7% (654)	11.8%	13.6%
	IIIb. Self-employed	4.9% (253)	4.4%	5.5%
	IIIc. Manual work supervisor	1.7% (87)	1.4%	2.1%
	IVa. Qualified manual worker	23.6% (1216)	22.5%	24.8%
	IVb. Manual worker	21.5% (1107)	20.4%	22.6%
	V. Unskilled worker	22.8% (1170)	21.7%	23.9%
Educational Level (11, 0.2%)	VI. Others	0.04% (2)	0%	0.2%
	Illiterate	2.1% (136)	1.8%	2.5%
	No formal education but can read and write	11.6% (756)	10.9%	12.4%
	Up to 5 years primary schooling (Early education)	19.5% (1263)	18.5%	20.4%
	Up to 8 years primary schooling	23.8% (1543)	22.7%	24.8%
	Up to 4 years secondary schooling (lower secondary)	9.1% (590)	8.4%	9.8%
	First cycle of vocational training	5.3% (345)	4.8%	5.9%
	Second cycle of vocational training	6.5% (420)	5.9%	7.1%
	Up to 6 years secondary schooling (upper secondary)	9.1% (593)	8.4%	9.8%
Employment Situation (3, 0.05%)	Short-cycle tertiary education (diploma) or Bachelor's degree	7.1% (465)	6.6%	7.8%
	Master's degree or PhD	5.9% (385)	5.4%	6.5%
	Employed	34.1% (2221)	33%	35.3%
	Unemployed but previously worked	21.5% (1398)	20.6%	22.5%
	Seeking first employment	1% (66)	0.8%	1.3%
	Retired (previously employed)	14.4% (938)	13.6%	15.3%
	Home keeper	18.1% (1178)	17.2%	19.1%
	Student	8.1% (527)	7.5%	8.8%
	Handicap/Permanent Disability	2.5% (160)	2.1%	2.9%
Economic difficulty to make ends meet (30, 0.5 %)	Other	0.2% (16)	0.2%	0.4%
	Very difficult	12.7% (819)	11.9%	13.5%
	Difficult	18.4% (1190)	17.5%	19.3%
	Somewhat difficult	30% (1944)	28.9%	31.1%
	Somewhat easy	25.5% (1651)	24.4%	26.5%
	Easy	12.5% (813)	11.8%	13.4%
Total net household income (1540, 23.7%)	Very easy	0.9% (60)	0.7%	1.2%
	Up to 300 euros	0.7% (36)	0.5%	1%
	From 301 to 499 euros	4.3% (212)	3.7%	4.9%
	From 500 to 999 euros	27.6% (1368)	26.3%	28.8%
	From 1000 to 1499 euros	37.6% (1866)	36.2%	38.9%
	From 1500 to 1999 euros	17% (844)	16%	18.1%
	From 2000 to 2499 euros	8.1% (401)	7.3%	8.9%
	From 2500 to 2999 euros	3.2% (157)	2.7%	3.7%
From 3000 to 4999 euros	1.5% (72)	1.2%	1.8%	
More than 5000 euros	0.2% (11)	0.1%	0.4%	

Table 9.3: Life habits of the study population (Andalusia, 2011).

Variables (missing responses; %)	Category (population aged +16; total n=6,507)	Percentage (n)	95% Confidence Interval	
			Minimum	Maximum
	Suspected alcoholism	3.1% (204)	2.7%	3.6%
Consumption of alcoholic beverages (2; 0.03%)	Yes, consumption of alcoholic beverage	44.4% (2891)	43.2%	45.6%
	Yes, but less than once a month	13.3% (864)	12.5%	14.1%
	No, no consumption of alcoholic beverage	42.3% (2750)	41.1%	43.5%
Smoker (3; 0.05%)	Yes, smoke daily	30.9% (2011)	29.8%	32.1%
	Yes, smoke but not daily	2.4% (157)	2.1%	2.8%
	Do not smoke but used to	17.5% (1137)	16.6%	18.4%
	Do not smoke or have never regularly smoked	49.2% (3199)	48%	50.4%
Body Mass Index [30] (0)	Low weight	3.8% (247)	3.4%	4.3%
	Normal weight	37.3% (2428)	36.1%	38.5%
	Overweight	39.7% (2585)	38.5%	40.9%
	Obesity I	17.1% (1113)	16.2%	18%
Physical activity in the workplace or usual activity (47; 0.7%)	Obesity II	2.1% (134)	1.7%	2.4%
	Sitting most of the workday	30.4% (1964)	29.3%	31.5%
	Standing most of the time without major movement or effort	55.8% (3605)	54.6%	57%
	Walking, carrying some weight. Frequent movement	11.5% (737)	10.7%	12.2%
Physical exercise in free time (3; 0.05%)	Hard work, tasks requiring major physical effort	2.4% (154)	2%	2.8%
	No physical activity	26.8% (1742)	25.7%	27.9%
	Occasional physical or sporting activity	55.9% (3639)	54.7%	57.1%
	Regular physical activity, several times a month	12% (779)	11.2%	12.8%
Dairy product consumption (7; 0.1%)	Physical training several times a week	5.3% (344)	4.8%	5.9%
	Daily	90.9% (5909)	90.2%	91.6%
	Three or more times a week	2.9% (191)	2.6%	3.4%
	One / two times a week	2.5% (163)	2.2%	2.9%
	Less than 1 time week	0.8% (52)	0.6%	1%
Fresh fruit consumption (11; 0.2%)	Never or almost never	2.8% (185)	2.5%	3.3%
	Daily	63.8% (4144)	62.6%	64.9%
	Three or more times a week	21.7% (1405)	20.7%	22.7%
	One / two times a week	9.6% (623)	8.9%	10.3%
	Less than once a week	2.9% (190)	2.5%	3.4%
Vegetables consumption (14; 0.2%)	Never or almost never	2.1% (134)	1.8%	2.4%
	Daily	41% (2665)	39.9%	42.1%
	Three or more times a week	34.1% (2211)	33%	35.2%
	One / two times a week	19.8% (1287)	18.9%	20.8%
	Less than once a week	3.7% (240)	3.3%	4.2%
Sweets consumption (30; 0.5%)	Never or almost never	1.4% (90)	1.1%	1.7%
	Daily	18.4(1191)	17.4%	19.3%
	Three or more times a week	25.2% (1629)	24.2%	26.2%
	One / two times a week	28% (1814)	27%	29.1%
	Less than once a week	14% (908)	13.2%	14.9%
	Never or almost never	14.4% (935)	13.6%	15.3%

older than those without CP (43.7 yrs; 95%CI=[43.2-44.2]; $p<0.001$). The most prevalent CP considered was Musculoskeletal Pain (MSP) (13.5%), followed by dysmenorrhea (4%), Chronic Cephalalgia (CC) (3.3%), and angina/chest pain (0.7%).

DISABLING CHRONIC PAIN: PREVALENCE BASED ONLY ON DESIGN WEIGHTS. Among the participants with CP, the pain was considered responsible for disability in some daily life activity (DCP) by 67%, i.e. 11.5% of the whole study population (95%CI=[10.73-12.28]; 793,455 individuals; Table 9.4), while nDCP was reported by 33%, 5.7% of the whole study population (95%CI=[5.2-6.3]) (393,080 individuals).

The most frequent disabling CP was MSP (72.3% of the population with MSP), followed by chest pain (69.7%), CC (61.5%), and dysmenorrhea (31.9%).

DCP was more frequent in females than in males (68.2% and 63.9%), although the difference did not reach statistical significance, and was significantly more frequent in the over-65-yr-olds (82.7%) than in other age groups. The mean age of the population with DCP was 58.5 yrs (95%CI=[57.2-59.8]), 13 yrs higher than that of the population with nDCP (45.3 yrs; 95%CI=[43.4-47.1]; $p<0.001$). DCP was more frequent from the age of 55 yrs onwards in both males and females ($p<0.001$). DCP was significantly more frequent in females versus males in all age groups ($p<0.001$), and this gender difference was greater in younger age groups. Thus, DCP prevalence was three-fold higher in females versus males up to the age of 45 yrs and two-fold higher at older ages. nDCP was significantly more frequent in females versus males up to the age of 54 yrs, with no significant gender difference at older ages. Finally, the absence of CP was significantly more frequent in males versus females in all age groups ($p<0.001$). No significant difference was observed between the mean age of males and females in the population with DCP (58.7 and 58.4 yrs, respectively; $p=0.8$), but a significant gender difference in mean age was found in those with nDCP (43.3 vs. 49.3 yrs, respectively, $p=0.035$).

DISABLING CHRONIC PAIN: CHARACTERISTICS. The prevalence of DCP was significantly higher ($p<0.001$) among: those who lived alone (19.5%), widows/widowers (29.6%), and unskilled workers (15.1%); those who were illiterate (28.8%), literate but with no schooling (24.9%), or only received primary schooling (15%); and those reporting difficulties to reach the end of the month (14.1%) and with a net household income < 1,000 €/month (17%). The prevalence of nDCP was not significantly associated with any of these characteristics.

Table 9.4: Prevalence of disabling chronic pain and other chronic diseases by sex and age (Andalusia, 2011)

Condition	Prevalence (n=6,507; men=3,209; women=3,298)		Disabling Condition (Subpopulation with the condition/s)		Disabling Condition (Total population)		
	Prevalence (Significance)	95% Confidence Interval	Prevalence (Significance)	95% Confidence Interval	Prevalence (Significance)	95% Confidence Interval	
Chronic Pain ^a	Total	17.2%	66.9%	64.14-69.55	11.5%	10.73-12.28	
	Women	23.9% (p<0.001)	(22,5-25,4)	68,2% (536)	(64,9-71,3)	16.3% (p<0.001)	15.1-17.6
	Men	10.3%	(9,3-11,4)	63,9% (212)	(58,6-68,8)	6.6%	5.8-7.5
	16-44 years	11.4%	10.4-12.5	48.2%	43.4-53.1	5.5%	4.8-6.3
	45-64	18.2%	16.5-20	70.4%	65.3-75	12.8%	11.4-14.4
	+65	32.3% (p<0.001)	29.7-35	82.7% (p<0.001)	78.6-86.1	26.7% (p<0.001)	24.2-29.2
	Total	45.9%	44.67- 47.08	59.7%	58-61.47	29.5%	28.45-30.66
	Women	52.9% (p<0.001)	51.2-54.6	64.5% (p<0.001)	62.2-66.7	36.8% (p<0.001)	35.2-38.5
	Men	38.6%	37-40.3	53.1%	50.3-55.8	22.1%	20.7-23.5
	16-44 years	23.7%	22.3-25.1	43.3%	40-46.7	13%	11.9-14.2
45-64	56.9%	54.6-59.1	56.6%	53.6-59.5	34%	31.9-36.2	
+65	92.1% (p<0.001)	90.4-93.5	74.9% (p<0.001)	72.3-77.4	69.7% (p<0.001)	67.1-72.3	
At least one Chronic Disease (besides Chronic Pain)							

a Chronic Pains: 'migraine/headache/chronic cephalalgia /frequent headache', 'angina/chest pain', 'back pain, neck pain, shoulder pain, waist pain, cervical/low back pain' or 'menstrual pain'.

As shown in [Table 9.5](#), a significantly higher likelihood of **DCP** was found in those belonging to manual social classes ($OR_{\text{manual}}=1.26$), those with lower schooling level $OR_{\text{illiterate/literate but with no schooling}} = 1.61$; $OR_{\text{Primary schooling}} = 1.57$), those who were unemployed and had worked previously versus those in employment ($OR = 1.33$), and residents of more rural areas ($OR = 1.28$).

DISABLING CHRONIC PAIN: OTHER CHRONIC DISEASES. A chronic disease other than **CP** was reported by 45.9% of the general population (3,169,331 individuals, 95% $CI=[3,086,051-3,252,610]$) ([Table 9.4](#)); the prevalence was 37% higher in females versus males, while the mean age (56.4; 95% $CI=[55.8-57]$) was slightly older than for **CP**. **DCP** prevalence was five-fold higher among those with other chronic diseases (20.4%) than in those without (20.4% vs. 3.9%, $p<0.001$), observing a similar result by sex and by age group [Table 9.6](#). The prevalence of **nDCP** was significantly higher in males with other chronic diseases versus those without (2-fold higher) and in the youngest age group versus older age groups (> 3-fold higher) [Table 9.6](#).

The prevalence of chronic diseases was three-fold higher in those with **DCP** than in those with **nDCP** ([Figure 9.2](#)). The ten most prevalent chronic diseases were the same in the different subpopulations, with the exception of prostate disorder, which was replaced with osteoporosis in the **DCP** subpopulation.

The prevalence of chronic depression or anxiety in the population with **DCP** was almost two-fold the prevalence in the population with **nDCP** and almost four-fold higher than in the population without **CP** (20.3%, 11.6%, and 5.8%, respectively; $p<0.001$). **DCP** prevalence was three-fold higher among those with depression or anxiety than among those without (29.8% vs. 10%; $p<0.001$), while this difference was smaller for the prevalence of **nDCP** (8.5% vs. 5.5%; $p<0.05$). The prevalence of depression or anxiety in the population with arterial hypertension and high cholesterol was significantly higher among those with **DCP** (23.3% and 30.9%, respectively) in comparison to those with no **CP** (11.9% and 13.2%) but not in comparison to those with **nDCP** (20.3% and 22.2%).

The prevalence of fibromyalgia was four-fold higher among those with **CP** than among those without (3.8% vs. 0.9%; $p<0.001$) and was 5% in the population with **DCP** versus 1.6% in the population with **nDCP** ($p<0.001$). Conversely, **DCP** prevalence was significantly (4-fold) higher in the population with fibromyalgia (39.7% vs. 11.1%; $p<0.001$), while the prevalence of **nDCP** was not (6.4% vs. 5.7%).

The prevalence of arthritis or rheumatism was almost four-fold higher among those with **CP** than among those without (22.3% vs. 5.8%; $p<0.001$) and was 28.4% in the population with **DCP** versus 9.9% in the popula-

Table 9.5: Factors associated with disabling chronic pain (a,b)

Disabling Chronic Pain (reference category: no Chronic Pain)	VARIABLES	Categories	p-value	Odds Ratio	95% Confidence Interval					
					Minimum	Maximum				
Yes	Social class (p=0.68)	I. Manager with 10<salaried staff	0.196	0.704	0.413	1.199				
		II. Manager with 10>salaried staff	0.231	0.776	0.512	1.176				
		IIIa. Administrative staff	0.215	0.806	0.573	1.134				
		IIIb. Self-employed	0.187	0.73	0.458	1.165				
		IIIc. Manual work supervisor	0.673	0.839	0.37	1.9				
		IVa. Qualified manual worker	0.603	0.929	0.704	1.226				
		IVb. Manual worker	0.836	0.973	0.748	1.265				
V. Unskilled worker		.	1	.	.					
No		Social class (p=0.68)	I. Manager with 10<salaried staff	0.321	0.68	0.317	1.458			
			II. Manager with 10>salaried staff	0.12	1.49	0.902	2.463			
			IIIa. Administrative staff	0.454	1.188	0.756	1.868			
			IIIb. Self-employed	0.977	0.991	0.52	1.886			
			IIIc. Manual work supervisor	0.232	0.295	0.04	2.187			
			IVa. Qualified manual worker	0.461	1.163	0.778	1.739			
	IVb. Manual worker		0.34	1.205	0.821	1.77				
V. Unskilled worker	.		1	.	.					
Yes	Social class (short version) (p=0.107)		Non-manual (I-III)	0.034	0.794	0.641	0.983			
No			Manual (IV-V)	.	1	.	.			
Yes	Level of Education (p=0.056)		Illiterate/No formal education	0.014	1.615	1.104	2.364			
			Primary education	0.008	1.57	1.127	2.187			
			Lower Secondary/First-cycle Vocational Training	0.139	1.366	0.903	2.066			
			Upper Secondary / Second-cycle Vocational Training	0.378	1.197	0.802	1.786			
		University education	.	1	.	.				
		No	Level of Education (p=0.056)	Illiterate/No formal education	0.347	1.268	0.773	2.081		
				Primary schooling	0.97	0.993	0.68	1.45		
Lower Secondary/ First-cycle Vocational Training				0.21	1.329	0.852	2.074			
Upper Secondary / Second-cycle Vocational Training				0.482	0.845	0.528	1.351			
University education				.	1	.	.			
Yes				Employment Situation (p<0.001)	Unemployed but previously worked	0.047	1.327	1.004	1.754	
					Seeking first job or student	0.031	0.444	0.213	0.929	
		Retired (previously employed)			<0.001	1.86	1.347	2.567		
		Home keeper			0.214	1.199	0.9	1.598		
	Handicap/Permanent Disability	<0.001			5.976	3.897	9.166			
	Employed	.			1	.	.			
	No	Employment Situation (p<0.001)			Unemployed but previously worked	0.943	0.988	0.714	1.368	
Seeking first job or student					0.552	0.848	0.492	1.46		
Retired (previously employed)					0.942	1.017	0.649	1.592		
Home keeper			0.121		0.737	0.502	1.084			
Handicap/Permanent Disability			0.217		1.587	0.763	3.303			
Employed			.		1	.	.			
Yes			Net monthly household income (p=0.024)		-999€	0.61	1.14	0.688	1.889	
	1000 - 1499 €				0.617	1.136	0.689	1.873		
	1500 - 2499€			0.448	0.817	0.484	1.378			
	+2500€			.	1	.	.			
	No			Net monthly household income (p=0.024)	-999€	0.606	0.84	0.432	1.631	
					1000 - 1499 €	0.449	1.28	0.676	2.424	
					1500 - 2499€	0.693	1.141	0.593	2.195	
+2500€					.	1	.	.		
Yes		Rurality Index (p=0.05)			Rurality (continuous)	0.02	1.28	1.04	1.576	
No					Rurality (continuous)	0.309	1.158	0.872	1.538	
Yes		Alcohol consumption (p=0.006)			Consumption of alcoholic beverages	0.865	0.983	0.806	1.199	
	No, less than once per month				0.05	1.272	1	1.619		
	No, no consumption of alcoholic beverages				.	1	.	.		
	No				Alcohol consumption (p=0.006)	Consumption of alcoholic beverages	0.132	0.813	0.621	1.064
			No, less than once per month			0.004	0.526	0.341	0.812	
			No, no consumption of alcoholic beverages			.	1	.	.	
			Yes			Tobacco consumption (p=0.023)	Smoke daily	0.001	1.449	1.165
Smoke but not daily							0.167	1.575	0.826	3.004
Do not smoke but used to				0.069			1.245	0.983	1.577	
Do not smoke or have never regularly smoked				.			1	.	.	
No	Tobacco consumption (p=0.023)			Smoke daily			0.854	1.027	0.773	1.364
				Smoke but not daily			0.629	0.797	0.318	2
				Do not smoke but used to			0.438	1.14	0.819	1.587
			Do not smoke or have never regularly smoked	.			1	.	.	
		Yes	Physical exercise in free time (p=0.2)	No exercise			0.565	1.187	0.662	2.129
				Occasional physical or sporting activity			0.75	1.098	0.619	1.947
				Regular physical activity, several times a month			0.422	0.762	0.392	1.48
Physical training several times to the week				.	1		.	.		
No				Physical exercise in free time (p=0.2)	No exercise		0.882	1.054	0.528	2.105
					Occasional physical or sporting activity		0.39	1.339	0.688	2.609
					Regular physical activity, several times a month	0.708	1.15	0.553	2.394	
		Physical training several times a week			.	1	.	.		
		Yes			Functional social support (p<0.001) (*Age, p=0.016)	Overall score	<0.001	0.969	0.953	0.984
						Overall score* age	0.006	1.00037	1.00011	1.00063
	Overall score					0.028	0.979	0.96	0.998	
Overall score* age	0.236					1.00022	0.99986	1.00058		
No	Functional social support (p<0.001) (*Age, p=0.016)					Yes	<0.001	0.343	0.192	0.612
						Yes* age	0.151	1.007	0.997	1.018
			Yes			0.008	0.341	0.154	0.752	
		Yes* age	0.012			1.021	1.005	1.039		

a Chronic Pains: 'migraine/headache/chronic cephalgia /frequent headache', 'angina/chest pain', 'back pain, neck pain, shoulder pain, waist pain, cervical/low back pain'.

b Variables with $p < 0.2$ are included in the multinomial logistic regression models (except for social class). Variables not included in the multivariate: missing data for over 3.5% of a subpopulation (e.g. working population, population over 65 yrs.), treated differently (e.g. categorization or coding), or replaced by a similar variable (e.g. SF-12 instead of self-rated health). All models were adjusted for age and sex. Significance level = 0.05. Interactions performed: sex * age, sex * independent variable analyzed and age * independent variable analyzed, showing the results that were statistically significant and did not affect the model convergence.

Table 9.6: Disabling chronic pain according to the presence of other chronic diseases

Chronic Pain ^a		POPULATION WITH OTHER CHRONIC DISEASES (total n=2987; n men=1,240; n women=1,747)		POPULATION WITHOUT OTHER CHRONIC DISEASES (total n=3,520; n men=1969; n women=1,551)	
		Percentage (Significance)	95% Confidence Interval	Percentage (Significance)	95% Confidence Interval
Disabling Chronic Pain	Total	20.4% (p < 0.001)	19-21.9	3.9%	3.3-4.6
	Women	25.7% (p<0.001)	23.7-27.8	5.7% (p=0.042)	4.7-7.0
	Men	13.1%	11.3-15.0	2.5%	1.9-3.3
	Ages 16-44	11.5%	9.5-13.8	3.0%	2.4-3.7
	45-64	19.3%	11.0-21.7	4.3%	3.2-5.9
	+65	28.1%(p < 0.001)	25.6-30.9	9.5%(p=0.001)	5.0-17.2
	Total	6.9% (p < 0.001)	6-7.9	4.7%	4-5.7
Non-disabling Chronic Pain	Women	8.0%	6.8-9.3	7.2% (p<0.001)	6.0-8.6
	Men	5.4% (p<0.001)	4.3-6.8	2.7%	2.1-3.5
	Ages 16-44	9.4%(p < 0.001)	7.5-11.6	3.0%	2.4-3.7
	45-64	6.4%	5.1-8.0	3.8%	2.7-5.4
	+65	5.5%	4.3-7.1	6.3%	2.9-13.3

a Chronic Pains: 'migraine/headache/chronic cephalalgia /frequent headache', 'angina/chest pain', 'back pain, neck pain, shoulder pain, waist pain, cervical/low back pain' or 'menstrual pain'.

tion with nDCP (p<0.001). Conversely, the prevalence of DCP was significantly (4-fold) higher in the population with arthritis or rheumatism (37.6% vs. 9%; p<0.001), but the prevalence of nDCP prevalence was not (6.5% vs. 5.6%).

Finally, the prevalence of osteoporosis was more than five-fold higher among those with CP than among those without (6.4% vs. 1.2%; p<0.001) and it was 8.9% in the population with DCP versus 1.4% in the population with nDCP (p<0.001). Conversely, DCP prevalence was significantly (almost 5-fold) higher in the population with osteoporosis versus the population without (49.4% vs. 10.7%; p<0.001), but the prevalence of nDCP was not (3.7% vs. 5.7%).

DISABLING CHRONIC PAIN: MULTIMORBIDITY. At least one chronic disease was present in 81.5% of the population with DCP Figure 9.3 versus 40.3% of the population without CP and 55.5% of the population with nDCP (p<0.001). At least three other chronic diseases were reported in 47.7% of the population with DCP versus 18.8% of the population with nDCP. There was a strong tendency for the frequency of multimorbidity to be higher in females versus males among those with DCP (83.4% and 76.4%; p=0.054) but not among those with nDCP (p=0.45). The mean number of chronic diseases in females with DCP was significantly higher than in the males with DCP (3.09, 95% CI= [2.85-3.33] vs. 2.32, 95% CI=[2.03-2.62]) and threefold higher than in females without CP (0.97, 95% CI =[0.92,1.03]).

With respect to the association of DCP with risk factors for chronic disease, the likelihood of DCP was higher in those who consumed alcohol less than once a month than in non-consumers. The likelihood of

Figure 9.2: Prevalence of chronic diseases (a) in subpopulations (b)

Chronic Pain Population		Non Chronic Pain Population		General Population		Disabling Chronic Pain Population		Non Disabling Chronic Pain Population	
Chronic Disease	%	Chronic Disease	%	Chronic Disease	%	Chronic Disease	%	Chronic Disease	%
Hypertension	34.71%	Hypertension	15.33%	Hypertension	18.66%	Hypertension	42.67%	Hypertension	18.61%
High cholesterol	23.78%	High cholesterol	9.94%	High cholesterol	12.32%	Arthritis / rheumatism	28.38%	High cholesterol	15.90%
Arthritis / rheumatism	22.27%	Diabetes*	6.71%	Arthritis / rheumatism	8.68%	High cholesterol	27.68%	Depression / Anxiety	11.64%
Poor blood circulation	18.52%	Arthritis / rheumatism	5.85%	Diabetes*	8.06%	Poor blood circulation	24.22%	Allergies*	10.52%
Depression / Anxiety	17.41%	Depression / Anxiety	5.82%	Depression / Anxiety	7.82%	Depression / Anxiety	20.26%	Arthritis / rheumatism	9.93%
Varicose veins in legs	16.00%	Allergies*	5.53%	Poor blood circulation	6.39%	Varicose veins in legs	19.66%	Varicose veins in legs	8.61%
Diabetes*	14.56%	Poor blood circulation	3.87%	Allergies*	6.25%	Diabetes*	18.42%	Poor blood circulation	7.01%
Allergies*	9.64%	Heart disorders*	3.15%	Varicose veins in legs	4.82%	Heart disorders*	10.30%	Diabetes*	6.76%
Heart disorders*	8.31%	Prostate disorders	3.09%	Heart disorders*	4.03%	Allergies*	9.20%	Prostate disorders**	5.83%
Prostate disorders**	6.94%	Varicose veins in legs	2.50%	Prostate disorders	3.48%	Osteoporosis	8.95%	Cataracts**	4.33%
Osteoporosis	6.43%	Asthma	2.08%	Asthma	2.32%	Prostate disorders**	7.56%	Heart disorders**	4.30%
Cataracts	6.16%	Cataracts	1.43%	Cataracts	2.24%	Cataracts	7.07%	Chronic constipation**	3.24%
Hard of hearing	4.63%	Lung*	1.41%	Osteoporosis	2.09%	Hard of hearing	6.26%	Asthma	3.23%
Chronic constipation	4.49%	Cancer*	1.36%	Hard of hearing	1.63%	Hemorrhoids	5.48%	Stomach problems*,**	2.95%
Hemorrhoids	4.30%	Heart attack*	1.24%	Lung*	1.63%	Urinary incontinence	5.23%	Colitis**,**	2.68%
Urinary incontinence	4.12%	Anemia*	1.19%	Anemia*	1.62%	Chronic constipation	5.10%	Lung**,**	1.90%
Fibromyalgia	3.84%	Osteoporosis	1.18%	Cancer*	1.54%	Fibromyalgia	4.95%	Hemorrhoids	1.90%
Anemia*	3.67%	Hard of hearing	1.01%	Heart attack*	1.53%	Anemia*	4.82%	Urinary incontinence**	1.87%
Asthma	3.49%	Kidney	1.00%	Fibromyalgia	1.43%	Kidney	4.28%	Heart attack**,**	1.61%
Stomach problems*	3.13%	Skin problems*	0.96%	Kidney	1.36%	Heart attack*	3.62%	Skin problems**,**	1.61%
Kidney	3.13%	Fibromyalgia	0.93%	Skin problems*	1.30%	Asthma	3.62%	Fibromyalgia	1.60%
Heart attack*	2.95%	Colitis*	0.91%	Colitis*	1.24%	Skin problems**	3.60%	Osteoporosis**	1.35%
Skin problems*	2.94%	Stomach problems*	0.78%	Hemorrhoids	1.21%	Stomach problems*,**	3.22%	Anemia**,**	1.35%
Colitis*	2.85%	Other mental	0.72%	Stomach problems*	1.18%	Stroke**,**	3.22%	Hard of hearing**	1.35%
Lung*	2.68%	Hemorrhoids	0.57%	Chronic constipation	1.17%	Lung**,**	3.07%	Cancer**,**	1.33%
Cancer*	2.40%	Stroke*	0.53%	Urinary incontinence	1.12%	Colitis**,**	2.93%	Other mental	0.81%
Stroke*	2.24%	Cirrhosis*	0.50%	Stroke*	0.83%	Cancer**,**	2.93%	Kidney**	0.80%
Cirrhosis**,**	1.43%	Urinary incontinence	0.50%	Other mental	0.72%	Cirrhosis**,**	1.86%	Infertility**	0.54%
Infertility**	0.71%	Chronic constipation	0.48%	Cirrhosis*	0.66%	Infertility**	0.80%	Cirrhosis**,**	0.54%
Other mental**	0.71%	Infertility**	0.14%	Infertility**	0.24%	Other mental	0.66%	Stroke**,**	0.26%

a Question: 'Has a healthcare professional [physician/nurse] told you that you suffer from ... ?'
 b Chronic Pains: 'migraine/headache/chronic cephalalgia /frequent headache', 'angina/chest pain', 'back pain, neck pain, shoulder pain, waist pain, cervical/low back pain' or 'menstrual pain'.
 * Cancer, malignant tumor, including leukemia, lymphoma; Diabetes / high blood sugar / sugar in urine; hypertension; Chronic colitis and intestinal diseases / inflammatory bowel disease / Cohn's disease; stomach ulcer / duodenum; Chronic lung disease / emphysema / chronic bronchitis; Cardiac disorders / cardiac failure / congestive cardiac failure; Myocardial infarction / heart attack; Chronic skin problems; Chronic allergies (such as rhinitis, eye inflammation, dermatitis, food allergies, etc.), excluding allergic asthma; Anemia or other blood disease; Stroke / cerebral hemorrhage; Hearing loss / hearing problems; Cirrhosis / hepatic disease / hepatic dysfunction;
 ** Coefficient of variation > 20%; therefore, the results should be interpreted with caution.

DCP was significantly higher in smokers (OR=1.45; $p=0.001$) and borderline significantly higher in ex-smokers (OR=1.24; $p=0.069$). Finally, the likelihood of DCP or nDCP was not significantly higher among those reporting lesser physical activity.

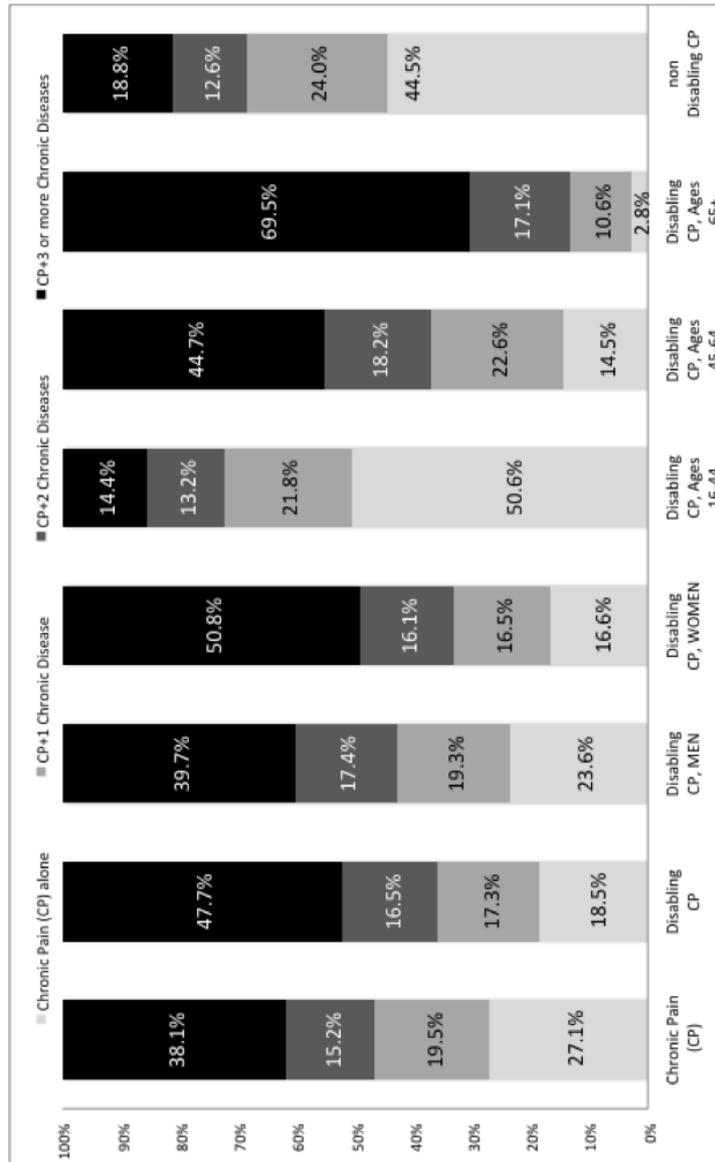
DISABLING CP: ASSOCIATED FACTORS. The final general multivariate model on factors associated with DCP (considering CC, MSP, and angina/chest pain; Figure 9.4 and Figure 9.5) was based on data from 96.65% of the study sample ($n=6289$) and was highly significant ($p<0.001$; $R^2_{Nagelkerke} = 0.27$). It included: gender ($p<0.001$); tobacco consumption ($p=0.074$); physical activity ($p=0.001$); physical limitations ($p=0.016$); hours of sleep (0.017); age ($p<0.001$); physical ($p<0.001$) and mental ($p<0.001$) quality of life, and their interaction with age ($p<0.001$ and $p=0.001$, respectively); environmental conditions ($p=0.002$); emotional social support ($p=0.028$); and number of chronic diseases ($p<0.001$).

Thus, the likelihood of DCP versus no CP was significantly higher in females (OR = 2.12, $p<0.001$), individuals sleeping ≤ 7 h (OR = 1.32, $p=0.004$), those with some physical limitation (OR = 1.61, $p=0.012$), and smokers (OR=1.42; $p=0.005$) but not significantly higher in those with 'heavy work, tasks requiring great physical effort' and ex-smokers; a higher likelihood of DCP (vs. no CP) was also observed with older age (OR_{10yrs} = 1.28; $p<0.001$), the presence of other chronic diseases (OR_{1additionalCD} = 1.26, $p<0.001$), worse environmental conditions (OR_{1-point decrease} = 1.16; $p=0.001$), worse physical (OR_{10-point decrease} = 2.38, $p<0.001$) or mental (OR_{10-point decrease} = 1.21, $p=0.001$) quality of life, and, although without reaching significance, lower emotional social support (OR_{10-point decrease} = 1.041, $p=0.096$). The likelihood of nDCP versus no CP was significantly higher in: females versus males (OR = 1.55, $p=0.001$); individuals with 'heavy work, tasks requiring great physical effort' versus those 'standing most of the time without much walking or major effort' (OR = 2.28, $p<0.001$) and those 'sitting during most of the day' (OR = 3.27, $p=0.009$); those with less emotional social support (OR_{10 point decrease} = 1.073, $p=0.023$); and those with other chronic diseases (OR_{1 additional CD} = 1.28, $p<0.001$). In contrast, the likelihood of nDCP was not significantly associated with physical or mental quality of life, age, environmental conditions, hours of sleep, physical limitations, or smoking.

DISABLING CHRONIC DISEASE: CALIBRATED PREVALENCE. The previous sections permitted the systematization and explanation of the calibration variable selection (see also Figure 9.1), which were finally sex, age group, and educational level.

Table 9.7 and Figure 9.6 show the bias, in terms of relative variations, of the direct population estimations for the calibration variables, with re-

Figure 9.3: Multimorbidity (a) and chronic pain (b)



a Question: 'Has a healthcare professional [physician/nurse] told you that you suffer from one or more of the following Chronic Diseases ... ?'

b Chronic Pains: 'migraine/headache/chronic cephalalgia /frequent headache', 'angina/chest pain', 'back pain, neck pain, shoulder pain, waist pain, cervical/low back pain' or 'menstrual pain'.

spect to their census values. Thus, the population estimations obtained from crossing sex with age group yielded deviations below 4% with respect to the population values observed in the INE (2015b). However, when educational level was included in the crossing, deviations were more than two-fold higher in the primary schooling or lower categories for the subpopulation aged between 16 and 64 yrs., being even much higher in women. This bias was the opposite for secondary schooling, or higher categories, increasing with the age groups, although to a lesser extent than primary schooling.

The population values of the Table 9.7 served to calculate new calibrated weights as a function of the procedure established by each calibration technique (item 2).

The logistic model adjusted for the three chronic pain categories (DCP, nDCP, non CP) with the calibration variables showed no statistically significant differences between the nDCP and non CP categories. Therefore, for the calibration of the prevalence of DCP according to the model-assisted approach, a binary logistic model was selected, considering DCP as dependent dichotomous variable (DCP category vs. nDCP or non CP). It was based on data from 99.8% of the study sample (n=6494), had a high significance ($p < 0.001$; $R^2_{Nagelkerke} = 0.15$) and included ($p < 0.001$) gender, age groups, and educational level. Second-order interactions were not statistically significant ($p > 0.45$).

Thus, the likelihood of DCP (vs. nDCP or non CP) was significantly higher in females (OR=2.56, $p < 0.001$) and individuals aged 45-64 yrs (OR=2.26, $p < 0.001$) or >65 yrs (OR=4.93, $p < 0.001$, in both categories versus individuals aged 16-44), and in those individuals with primary schooling or lower educational level vs. university education (OR=1.72, $p = 0.001$).

Table 9.8 and Figure 9.7 depict the statistical results of DCP according to the different calibration approaches. All calibration techniques, compared to the estimation based on sampling design weights, produced lower prevalence, population, and variance estimations. Thus, the prevalence of DCP was 11.51% with the original sampling weights, 10.76% with linear calibration based on marginals, 10.91% with linear calibration based on crossings, and 10.78% with model-assisted calibration. All Coefficient of Variation (CV) were below 4%, guaranteeing high precision for all estimations.

CHANGES AFTER CALIBRATING. All DCP prevalence calibrations, depicted in Figure 9.8, reduced by more than 5.2% all the direct estimations for the whole sample. Regarding the population estimations, the marginal linear calibration and the model-assisted calibration decreased the estimated population with DCP by more than 50,000 individuals. Fi-

Table 9.7: Observed biases of the calibration variables

Age	Educational Level	Men, estimated population ^a	Women, estimated population ^a	Total, estimated population ^a	Men ^b	Women ^b	Total ^b	Men bias ^c	Women bias ^c	Total bias ^c
16-44	Primary schooling or lower	709,002	715,448	1,424,449	307,360	217,325	524,685	130.7%	229.2%	171.5%
	Secondary schooling	852,243	718,123	1,570,366	1,192,015	1,094,065	2,286,080	-28.5%	-34.4%	-31.3%
	University or higher	307,463	328,202	635,665	305,110	420,470	725,580	0.8%	-21.9%	-12.4%
	Total	1,868,707	1,761,773	3,630,480	1,804,485	1,731,860	3,536,345	3.6%	1.7%	2.7%
45-64	Primary schooling or lower	615,963	730,795	1,346,757	316,005	359,070	675,075	94.9%	103.5%	99.5%
	Secondary schooling	245,974	181,773	427,747	540,110	521,950	1,062,060	-54.5%	-65.2%	-59.7%
	University or higher	134,576	89,842	224,418	172,950	162,490	335,440	-22.2%	-44.7%	-33.1%
	Total	996,512	1,002,410	1,998,922	1,029,065	1,043,510	2,072,575	-3.2%	-3.9%	-3.6%
>65	Primary schooling or lower	464,744	690,893	1,155,637	351,850	549,670	901,520	32.1%	25.7%	28.2%
	Secondary schooling	47,721	23,258	70,978	138,035	132,415	270,450	-65.4%	-82.4%	-73.8%
	University or higher	22,380	19,034	41,414	55,190	33,720	88,910	-59.4%	-43.6%	-53.4%
	Total	534,845	733,185	1,268,029	545,075	715,805	1,260,880	-1.9%	2.4%	0.6%
TOTAL		3,400,064	3,497,368	6,897,432	3,378,625	3,491,175	6,869,800	0.6%	0.2%	0.4%

^a Based on sample design (original sampling weights).

^b According to INE (2015b).

^c Percentage difference between the population estimation and the population according to INE (2015b).

Table 9.8: Calibrated prevalence and population with disabling chronic pain according to different approaches

Disabling Chronic Pain			Sampling design weights	Marginal linear calibration	Crossing linear calibration	Assisted-model calibration (linear)
Prevalence (%)	Estimation		11.51	10.76	10.91	10.78
	95% Confidence Interval	Minimum	10.73	10.01	10.16	10.03
		Maximum	12.28	11.51	11.65	11.52
	Sampling error		0.39616	0.38084	0.38089	0.38082
Population	Estimation		793,455	739,093	749,324	740,266
	95% Confidence Interval	Minimum	739,915	687,813	698,038	688,989
		Maximum	846,995	790,372	800,610	791,542
	Sampling error		27,316	26,163	26,166	26,162

nally, model-assisted calibration achieved the greatest variance reduction (8.28%), almost with no differences compare to the rest of calibration approaches.

9.4 DISCUSSION

The definition of CP in this study was the medical or healthcare professional diagnosis, following the criteria of the Andalusian Health Strategy for the care of people in pain (Álvarez-González et al., 2010) and the International Association for the Study of Pain (Álvarez-González et al., 2014). The CP prevalence among adults (over 16 yrs old) residing in Andalusia was 17.2%, comparable to the prevalence among over-18-yr-olds of 18.34% reported by Langley et al. (2011) in the same region and of 17.25% and 16.6% published by other studies in Spain (Azevedo et al., 2012). It is within the range of 12 to 30% reported by Breivik et al. (2006) for moderate and severe CP among over-18-yr-olds in Europe, including Spain (12%). The most prevalent type of CP was MSP (13.5%), followed by CC (3.3%), in agreement with previous reports (Álvarez-González et al., 2010; Breivik et al., 2006; Johannes et al., 2010; Langley et al., 2011). The mean age of the present population was 54.1 yrs, within the reported range in Spain of 50 (Johannes et al., 2010; Langley et al., 2011) to 56.5 yrs (Dueñas et al., 2014). The mean age was 58.5 yrs in the DCP population, 45.3 yrs in the CP population, and 43.7 yrs in the population without CP. There was a negligible difference in mean age between those with nDCP and those without CP. DCP affects two-thirds of the population of Andalusia with CP (11.5% of the general population), largely affecting females (16.3% of the general female population) (Bingefors and Isacson, 2004; Blyth, March, AJ Brnabic, et al., 2006) and

over-65-yr-olds (26.7% of the general elderly population) (Herrera-Silva et al., 2012; Soares-Weiser and Sola, 2014). Almost half of the population aged over 65 yrs with CP required some instrumental help and more than one third needed basic help, and these proportions were significantly higher in the population with DCP (Björnsdóttir et al., 2013; Kennedy et al., 2014).

As mentioned above, females are more frequently affected by DCP in comparison to males and this gender difference was greater in the lower age groups in our study (Blyth, March, AJ Brnabic, et al., 2006). In comparison to nDCP, DCP was more closely associated with low socioeconomic level (income <1500 €/month), receipt of disability pension, and health-related unemployed status (Blyth, March, AJ Brnabic, et al., 2006). In our study, the presence of DCP (adjusted for sex and age) was also associated with having primary education, being unemployed after having worked previously, unqualified/unskilled employment, rural life, and smoking.

CP is also associated with worse health-related quality of life (Álvarez-González et al., 2014; Breivik et al., 2006; Dueñas et al., 2014; Johannes et al., 2010; Langley et al., 2011). Langley et al. (2011) observed a moderate reduction (21%) in the mental component and a large reduction (2-3-fold) in the physical component among individuals with CP. Similar associations were observed in the populations with DCP and nDCP, but statistical significance was not reached in the latter. The mean score for functional social support (affective and confidant components), considered to play an important role in helping sufferers cope with their pain (Gibson and Lussier, 2012; Leonardi et al., 2006), was significantly lower in those with CP than in those without ($p < 0.01$). Again, the population with DCP obtained a significantly lower score (vs. no CP) in both components, whereas the score was only significantly lower for those with non-disabling CP versus no CP in the affective component.

Other chronic diseases were reported by almost three-quarters (72.9%) of the population with CP, by less than half (40.3%) of the population without CP, and by over four-fifths (81.5%) of the population with DCP. In the population with DCP, multimorbidity was more frequent in females versus males, in over-65-yr-olds, and in those with low economic level (Kerns et al., 2002); these differences were not observed in the population with non-disabling CP. Thus, 50.8% of females versus 39.7% of males had three or more chronic diseases in the DCP population, while 20.8% of females and 17.9% of males had three or more chronic diseases in the non-disabling CP population. This gender difference was greater with higher age in the population with DCP. In general, the prevalence of chronic pain was more than three-fold higher and the prevalence of DCP five-fold higher among those with chronic diseases than among

those without. The gender difference in the prevalence of DCP was even greater among those with other chronic diseases. According to allostatic load models (McEwen, 1998), CP is more disabling in patients with a larger number of chronic diseases, increasing the health risk (Dominick et al., 2014; Eriksen et al., 2003). In comparison to the population without CP, the prevalence of diseases such as fibromyalgia, arthritis, or rheumatism/osteoporosis was significantly higher among those with DCP and was only non-significantly higher among those with nDCP.

Cardiovascular diseases have been associated with CP (Evers et al., 2003). In the present study, arterial hypertension and high cholesterol were more frequent among those with CP than among those without, as previously reported (Evers et al., 2003; Fox et al., 2006), supporting reports of a higher risk of hypertension in individuals with CP (Ohayon and Stingl, 2012). The prevalence of arterial hypertension in the DCP population was more than double that in non-disabling CP or that in no CP. The mechanisms underlying the association between CP and hypertension have not been fully elucidated, and the allostatic factors involved remain under discussion (Fox et al., 2006). There has been considerable research on the association between CP and depression or anxiety (Asmundson and Katz, 2009; Bair et al., 2008; Baune et al., 2008; Bruehl et al., 2005; Ohayon and Schatzberg, 2010; Olsen et al., 2013; Tsang et al., 2008), with reports of CP in up to 60% of individuals with depression (Evers et al., 2003). In the present study, depression or anxiety was recorded in 17.4% of those with CP and 5.85% of those without, slightly higher than the percentages of 15% and 3%, respectively, reported by a study in Ireland that only considered depressive symptoms (Raftery et al., 2011). The prevalence of depression or anxiety in the present population with DCP was two-fold higher than in the population with non-disabling CP and four-fold higher than that without CP. Conversely, the prevalence of DCP was three-fold higher in the population with depression or anxiety than in the population without, signifying an increase in disability when CP is associated with depression or anxiety (Azevedo et al., 2012; Baune et al., 2008; Blyth, March, AJ Brnabic, et al., 2006; Raftery et al., 2011; Reid et al., 2015; Tsang et al., 2008).

One study strength is the source of data (EAS), a very large population survey based on face-to-face interviews more than 6,500 adults. This sample size allows precise and reliable estimations to be made. Besides information on CP, this survey gathers a large amount of information on other diseases, activity limitations, general and employment health, and utilization of healthcare services, among other characteristics of the population, allowing comprehensive analysis of CP and associated factors. Survey limitations detected in the present study include

the need to add muscle and joint pain in the lower and upper extremities (except shoulder) and various traumatological, post-surgical, and neuropathic conditions and the need to separate data on depression and anxiety. It would also be preferable to gather direct data on CP with a simple question (Dominick et al., 2014), avoiding its construction based on other chronic diseases and reducing the possibility of overestimating its prevalence. It is also recommended to gather data on the time since CP onset, using 6 months as the criterion for chronicity (Breivik et al., 2006; Gerrits et al., 2015; Portenoy et al., 2004)(Breivik et al., 2006), and on its intensity [58]. All of these changes have been proposed for the 2015/16 edition of the EAS, enabling analysis of possible biases (Von Korff, 1992).

A further strength of the study is the calibration of DCP prevalence and population estimations. In its absence, the estimations would have been overestimated and less accurate because of major educational level differences between the sample and the study population (or sampling framework; Table 9.7). With respect to the different calibration approaches used, only the results for the linear distance are presented in this study, because those for the raking distance were the same.

Marginal calibration (approach 2) permits the incorporation of more auxiliary variables in the construction of the new sampling weights, because it only requires the total marginal population for each auxiliary variable (A. Cabrera-León et al., 2015), while the crossing and model-assisted methods require the total marginal populations for each sub-population resulting from crossing auxiliary variables. It is also necessary to have at least two sample elements in each crossing and, in the case of model-assisted calibration, also in the crossing with the study variable. Both requirements can markedly limit crossing calibration or model-assisted techniques, especially when the study population, sample, or study variable prevalence is reduced. Thus, if the study population is small, there may be confidentiality concerns with regard to the population record from which auxiliary variables are extracted. On the other hand, if the sample or prevalence is low, there can be problems of an inadequate sample for certain crossing categories.

In the present study, three auxiliary variables (sex, age, and education) met the calibration conditions Figure 9.1. The ability to include more auxiliary variables would not necessarily yield higher estimation precision. In fact, for marginal calibration, it was possible to include auxiliary variable employment status, excluded in the other two calibration approaches due to insufficient samples in the crossings, and a slightly higher variance was found than that obtained for calibration with three auxiliary variables.

Another limitation of model-assisted calibration is that the new weights obtained could only be applied in estimations for the study variable, because they were calculated according to the probabilities obtained from the adjusted model. However, the remaining calibration techniques would serve for other study variables, with the only condition being a significant association with the auxiliary variables (hence, the requirement in previous studies for bivariate and multivariate association between the study and auxiliary variables). Thus, these calibrated inferences are entirely design-based. A set of calibrated weights is computed by modifying the design weights through the available auxiliary information. These weights are used to compute the (nearly design-unbiased) calibration estimator for any study variable total in a multi-purpose survey. The observed study variable values do not affect the calibrated weights. This is an important feature of calibration estimation (VM Estevao and Särndal, 2006).

Regardless of these limitations, model-assisted calibration was the best technique to estimate DCP prevalence, although it showed almost no differences with estimation obtained with the other calibration approaches.

In addition to the above requirements or limitations, the calculation of the new calibrated sample weights requires that there are no important sample losses in calibration variables, and there should be no significant differences between valid and lost samples. In our study, 7 of the 6,505 participants were lost due to calibration variables, leaving 99.9% of the total sample.

There was no non-response problem in our study, with only four missing values for DCP, the main study variable. Nevertheless, non-response is a frequent difficulty in surveys, and calibration is a useful approach to meet this challenge. The new weights computed are, at least for most elements, greater than the weights that would have been applied in the case of full response. This is in order to compensate for elements lost due to nonresponse (Särndal and Lundström, 2005). Calibration can also be used both to adjust the nonresponse and to increase the statistical efficiency of estimators (PS Kott and Day, 2014).

An alternative technique to adjust weightings is the pseudo-empirical likelihood method (PEL, J. Chen and R Sitter 1999). Although calibration and PEL methods are asymptotically equivalent, calibration is much easier to explain to users than PEL.

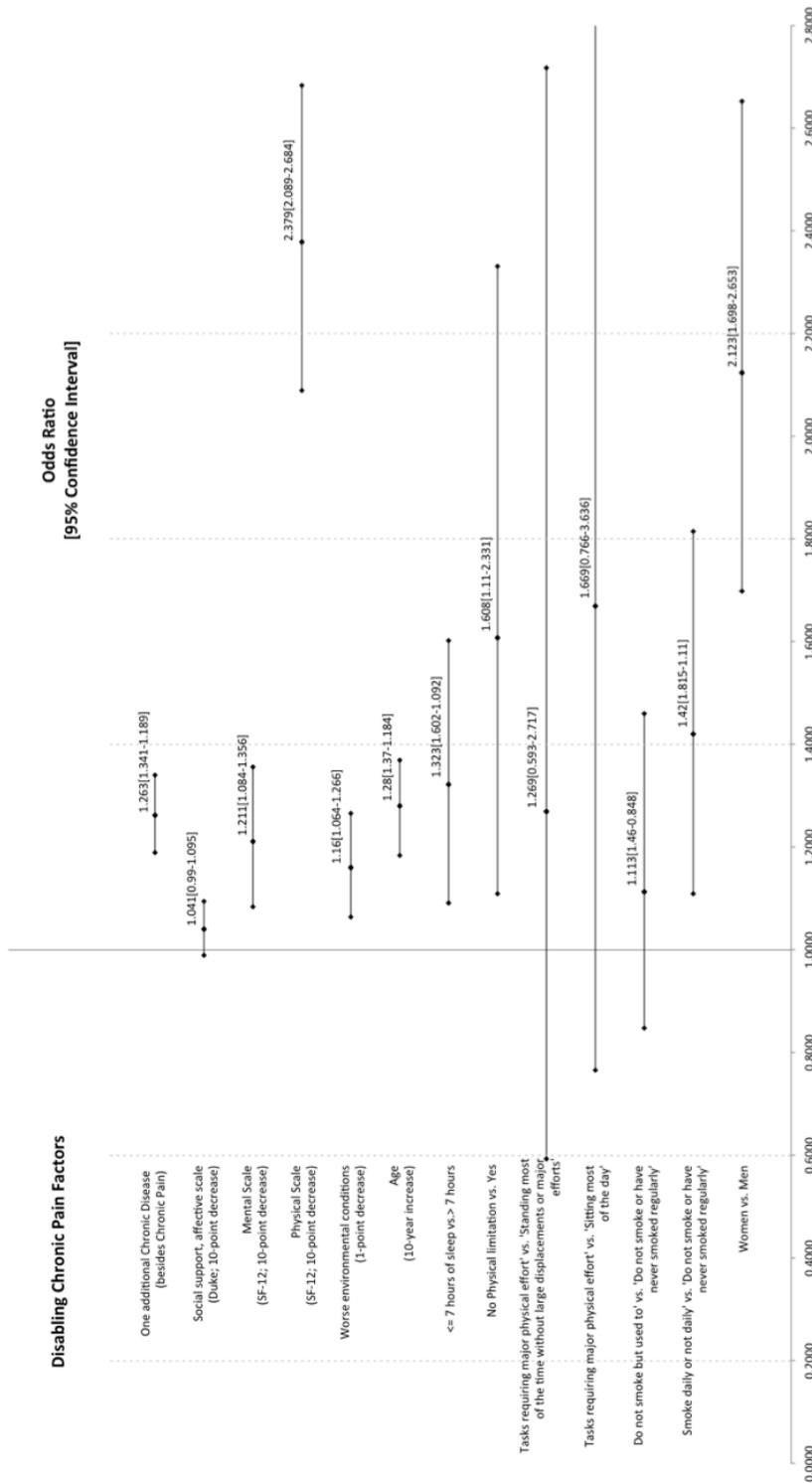
In summary, disabling CP is highly frequent (66.9%) among adults with chronic pain in Andalusia and relatively frequent (11.5% without calibration and in the general adult population). Disabling CP was significantly associated with elevated multimorbidity, disfavored or vulnerable social status (e.g., female sex, advanced age, unemployment, un-

skilled employment, low income, and low affective social support), poor health habits, worse environmental or work conditions, and impaired quality of life. In addition, the population with disabling CP showed statistically significant differences in the majority of study variables with the population without CP and even with the population with non-disabling CP. Indeed, very few statistically significant differences were found between those without CP and those with non-disabling CP. The present results contribute evidence on the need to construct and apply simple tools for the identification of individuals with DCP and to focus future research efforts and healthcare and social interventions on this population.

This article also offers a proposal to measure and assess disabling chronic pain, which may be useful for cross-disciplinary comparisons of chronic pain and disability in clinical and population-based studies. They may also assist health managers to tailor health resources and services to the real needs of their populations and provide a more rapid and successful response to individuals who suffer, or even those who are going to suffer, chronic pain and its disabling effect.

Finally, the calibration adjustments applied to the DCP estimations increased the validity of their generalization and also improved their accuracy. While linear calibration has already been applied in others population-based health survey by the authors of this study (A. Cabrera-León et al., 2015) and others (Legleye et al., 2015), this is the first time that a calibration technique has been applied to the Andalusian Health Survey. Furthermore, to the best of our knowledge, this is the first time that model-assisted calibration has been applied to any population-based health survey. These calibration approaches are being increasingly adopted in social sciences, especially in studies with small samples or with coverage or non-response problems. The methodology developed in our study may serve as a model for application in other epidemiological studies.

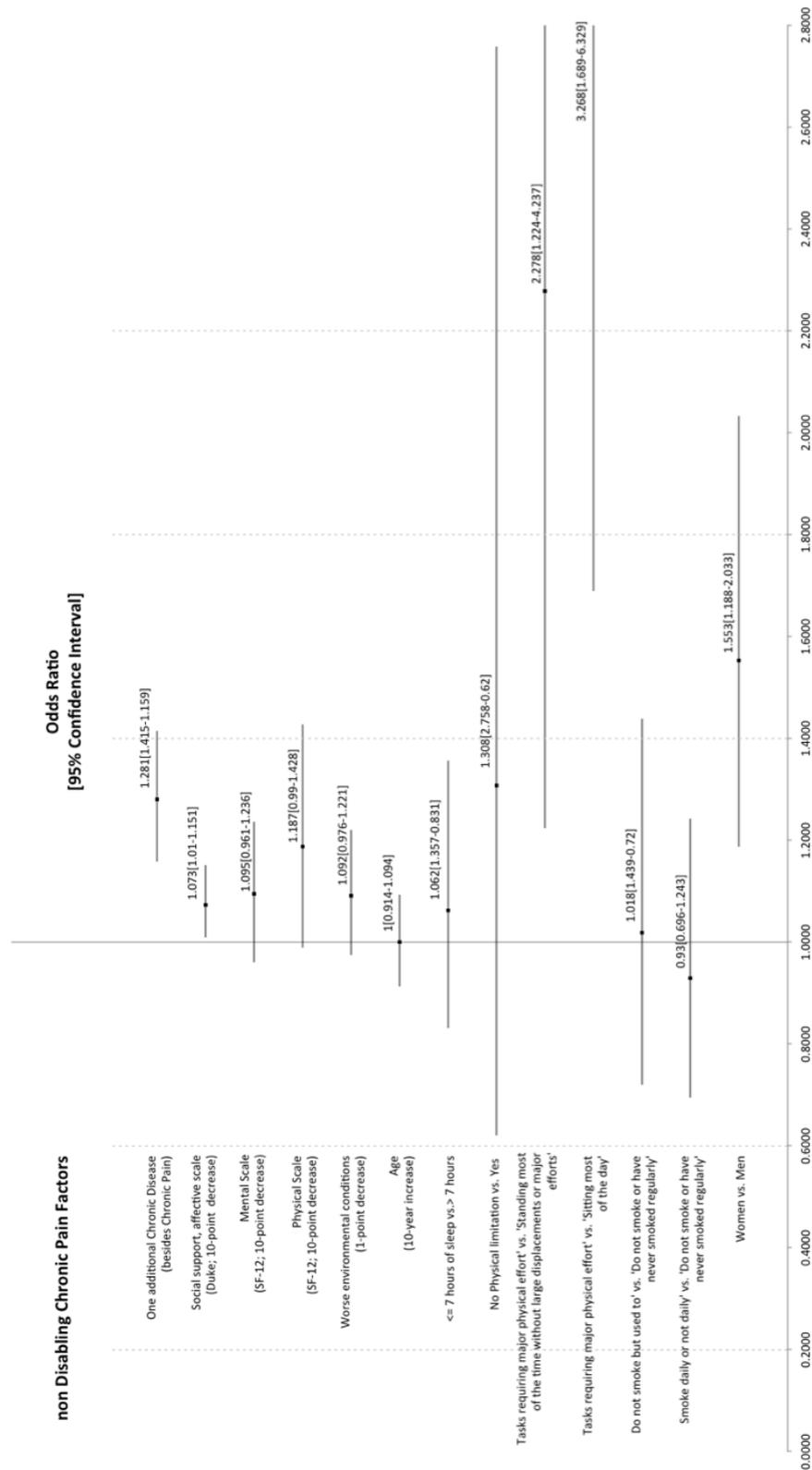
Figure 9.4: Factors associated with disabling chronic pain (a,b)



a Chronic Pains: 'migraine/headache/chronic cephalalgia /frequent headache', 'angina/chest pain', 'back pain, neck pain, shoulder pain, waist pain, cervical/low back pain'.

b The odds ratios are calculated based on the following reference categories/values: Sex = Male; Tobacco = Do not smoke or have never smoked regularly; Physical activity = Tasks requiring major physical effort; Physical limitations = Yes; Sleep > 7 hours; Age = 45.4; Physical Component Score (PCS_SF-12) = 51.12; Mental Component Score (MCS_SF-12) = 50.67; Affective social support (Duke) = 81.01; Number of chronic diseases (besides CP) = 1.07; Environmental conditions = 5.83.

Figure 9.5: Factors associated with non-disabling chronic pain (a,b)



a Chronic Pains: 'migraine/headache/chronic cephalalgia /frequent headache', 'angina/chest pain', 'back pain, neck pain, shoulder pain, waist pain, cervical/low back pain'.

b The odds ratios are calculated based on the following reference categories/values: Sex = Male; Tobacco = Do not smoke or have never smoked regularly; Physical activity = Tasks requiring major physical effort; Physical limitations = Yes; Sleep > 7 hours; Age = 45.4; Physical Component Score (PCS_SF-12) = 51.12; Mental Component Score (MCS_SF-12) = 50.67; Affective social support (Duke) = 81.01; Number of chronic diseases (besides CP) = 1.07; Environmental conditions = 5.83.

Figura 9.6: Observed biases (calibration variables)

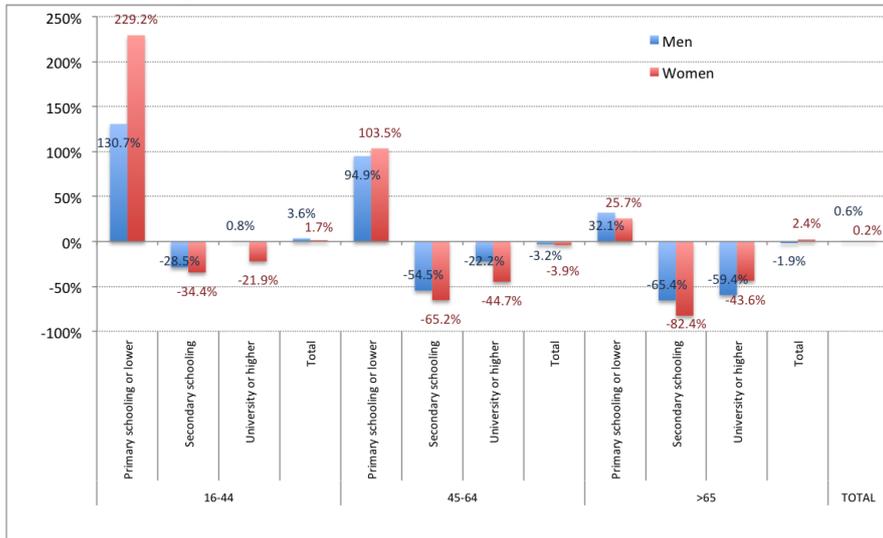


Figure 9.7: Calibrated prevalence and population with disabling chronic pain according to different approaches

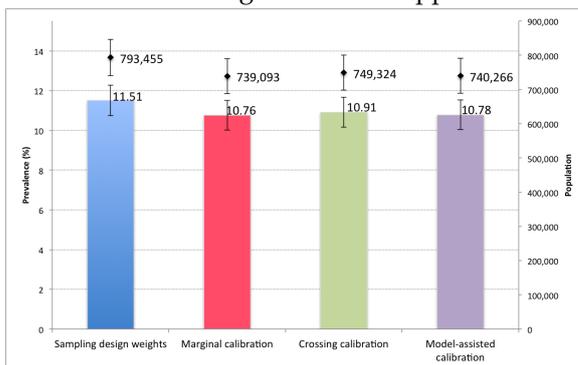
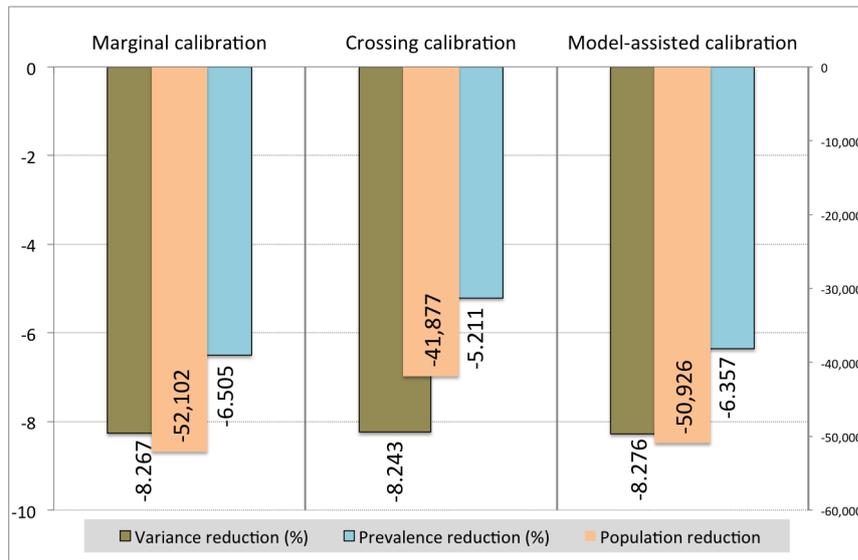


Figure 9.8: Percentage reductions in variance and prevalence estimations after calibration



Parte V

APÉNDICE

Este apéndice incluye el código utilizado en R para calibrar las prevalencias de esterilidad ([chapter 8](#)) y del dolor crónico discapacitante (DCP, [chapter 9](#)).

A.1 PREVALENCIA CALIBRADA DE ESTERILIDAD

El [Listing A.1](#) prepara R, instalando los paquetes necesarios, cargando la base de datos definiendo las variables dependientes y los pesos muestrales de diseño.

[Listing A.2](#) permite obtener las estimaciones con los pesos originales del diseño para la proporción (prevalencia) y total de DCP así como para su CI, error muestral y CV. Se muestra el código para una de las variables dependientes (esterilidad subjetiva), replicando el código para el resto de variables.

En [Listing A.3](#) se obtienen las estimaciones de calibración para España mediante el método lineal y según los siguientes vectores de variables auxiliares: edad; y edad y estado civil.

En [Listing A.4](#) se obtienen las estimaciones de calibración para España mediante el método lineal y según el vector de variables auxiliares de edad, estado civil y situación laboral.

En [Listing A.5](#) se obtienen las estimaciones de calibración para España mediante el método lineal y según los siguientes vectores de variables auxiliares: edad, estado civil y situación laboral; edad, estado civil, situación laboral y nivel de estudios.

En [Listing A.6](#) se continúa con las estimaciones de calibración para España mediante el método lineal y, en esta ocasión, según el vector de variables auxiliares de edad, estado civil, situación laboral, nivel de estudios y país de nacimiento.

En [Listing A.7](#) se continúa con las estimaciones de calibración para España mediante el método lineal y, en esta ocasión, según el vector de variables auxiliares de edad, estado civil, situación laboral, nivel de estudios, país de nacimiento y número de hijos.

[Listing A.8](#) sirve para obtener una tabla con todas las estimaciones para una variable dependiente (esterilidad subjetiva) según cada uno de los vectores anteriores de variables auxiliares.

Con el [Listing A.9](#) permite extraer las variaciones porcentuales de la calibración lineal según los distintos vectores de variables auxiliares con respecto a la estimación con los pesos originales de diseño.

Finalmente, para cada variable dependiente de esterilidad, se escogieron aquellos vectores de variables auxiliares que presentaron una mayor reducción de la varianza (véase [Figure 8.4](#)).

A.2 PREVALENCIA CALIBRADA DE DOLOR CRÓNICO DISCAPACITANTE

El [Listing A.10](#) prepara R, instalando los paquetes necesarios, cargando la base de datos y definiendo las variables así como los pesos muestrales de diseño.

[Listing A.11](#) permite obtener las estimaciones con los pesos originales del diseño para la proporción (prevalencia) y total de DCP así como para su CI, error muestral y CV. Se muestran igualmente los resultados obtenidos.

En [Listing A.12](#) se obtienen las estimaciones según calibración lineal para las sumas marginales de las variables auxiliares.

En [Listing A.13](#) se obtienen las estimaciones según calibración lineal para las sumas de sus cruces de las variables auxiliares.

[Listing A.14](#) muestra cómo calcular las estimaciones según pesos calibrados por el método raking y, al igual que con el método lineal, para las sumas marginales de las variables auxiliares y para las sumas de sus cruces.

[Listing A.15](#) muestra cómo estimar el DCP según calibración modelo asistida para el método lineal y raking.

[Listing A.16](#) muestra cómo estimar el DCP según calibración modelo asistida para el método raking.

[Listing A.17](#) sirve para obtener una tabla con todas las estimaciones anteriores.

Con el [Listing A.18](#) permite extraer las variaciones porcentuales de los distintos métodos de calibración con respecto a la estimación con los pesos originales de diseño

Listing A.1: Infertility: starting R

```
### PREPARACIÓN R ###
setwd("...")
read.table("datosesterilidad4.csv",head=T,sep=";", dec=",")->data
library("sampling",lib.loc="/Library/Frameworks/R.framework/Versions
/3.1/Resources/library")
library("samplingVarEst",lib.loc="/Library/Frameworks/R.framework/
Versions/3.1/Resources/library")

###Definiendo las variables dependientes###
data$uno->uno
data$Esterilidad_diagnosticada->Ydiag
data$Esterilidad_subjetiva->Ysubj
data$Esterilidad_12meses->Yeste12
data$Esterilidad_primaria->Yprim
data$Esterilidad_secundaria->Ysec
data$Subfertilidad_6->Ysub6
data$Subfertilidad_12->Ysub12
data$Subfertilidad_24->Ysub24

#Estableciendo las variables con valores válidos para estimaciones
sobre la muestra válida
Ydiag!=3->Ydiagc3
Ysubj!=3->Ysubjc3
Yeste12!=3 & Yeste12!=4 & Yeste12!=9->Yeste12c349
Yprim!=3 & Yprim!=8 & Yprim!=9->Yprimc389
Ysec!=3 & Ysec!=8 & Ysec!=9->Ysecc389
Ysub6!=8->Ysub6c8
Ysub12!=8->Ysub12c8
Ysub24!=8->Ysub24c8

#definiendo los pesos
data$peso->wi
pik<-1/wi
```

Listing A.2: Estimation based on sampling design weights

```

#PREVALENCIAS NO CALIBRADAS DE MUESTRA NO PERDIDA Y SUBPOBLACIONES
####
#Se muestra el código para una de las variables dependientes,
  replicando el código para el resto de variables
### ESTIMADORES: PROPORCIÓN, IC PROPORCIÓN, VARIANZA, ERROR MUESTRAL
  , CV, TOTAL, IC TOTAL, LONGITUD TOTAL.

sum(Ysubj[Ysubj3]*wi[Ysubj3])->TYsubj  varest(Ys=Ysubj[Ysubj3],Xs=
  NULL,pik[Ysubj3],w=wi[Ysubj3])->VTYsubj  round(100*c(TYsubj
  -1.96*sqrt(VTYsubj),TYsubj+1.96*sqrt(VTYsubj))/sum(wi[Ysubj3])
  ,5)->ICYsubj
round(c(TYsubj-1.96*sqrt(VTYsubj),TYsubj+1.96*sqrt(VTYsubj)),5)->
  ICYsubjTOT  round((TYsubj+1.96*sqrt(VTYsubj))-(TYsubj-1.96*sqrt(
  VTYsubj)),5)->ICYsubjTODIF  round(c(100*TYsubj/sum(wi[Ysubj3]),
  ICYsubj, VTYsubj, 100*sqrt(VTYsubj)/sum(wi[Ysubj3]), 100*(sqrt
  (VTYsubj)/sum(wi[Ysubj3]))/(TYsubj/sum(wi[Ysubj3])), TYsubj,
  ICYsubjTOT, ICYsubjTODIF),5)->resYsubj

#PREVALENCIAS NO CALIBRADAS DE TODA LA MUESTRA (PERDIDOS Y TODA LA
  POBLACIÓN)####

round(100*c(TYsubj-1.96*sqrt(VTYsubj),TYsubj+1.96*sqrt(VTYsubj))/sum
  (wi),5)->ICYsubjPOB
round(c(100*TYsubj/sum(wi), ICYsubjPOB, VTYsubj, 100*sqrt(VTYsubj)/
  sum(wi), 100*(sqrt(VTYsubj)/sum(wi))/(TYsubj/sum(wi)), TYsubj,
  ICYsubjTOT, ICYsubjTODIF),5)-> resYsubjPOB

```

Listing A.3: Linear calibration (age, marital status)

```

#CALIBRACIÓN POR EDAD ESPAÑA####
TedadESP<-c(1894440,1980630,1874620,1785945)

calib(Xs=Xedad,d=wi,total=TedadESP,method="linear")->gpESP

sum(Ysubj[Ysubj3]*gpESP[Ysubj3]*wi[Ysubj3])->TYsubjcalESP
varest(Ys=Ysubj[Ysubj3],Xs=Xedad[Ysubj3,],pik[Ysubj3],w=wi[
  Ysubj3]*gpESP[Ysubj3])->VTYsubjcalESP
round(100*(TYsubjcalESP-1.96*sqrt(VTYsubjcalESP),TYsubjcalESP+1.96*
  sqrt(VTYsubjcalESP))/sum(TedadESP),8)->ICYsubjcalESP
round(c(TYsubjcalESP-1.96*sqrt(VTYsubjcalESP),TYsubjcalESP+1.96*sqrt
  (VTYsubjcalESP)),5)->ICYsubjTOTcalESP
round((TYsubjcalESP+1.96*sqrt(VTYsubjcalESP))-(TYsubjcalESP-1.96*
  sqrt(VTYsubjcalESP)),5)->ICYsubjTOTDIFcalESP
round(c(100*TYsubjcalESP/sum(TedadESP), ICYsubjcalESP, VTYsubjcalESP
  , 100* sqrt(VTYsubjcalESP) / sum(TedadESP), 100*( sqrt(
  VTYsubjcalESP)/sum(TedadESP))/ (TYsubjcalESP/sum(TedadESP)),
  TYsubjcalESP, ICYsubjTOTcalESP, ICYsubjTOTDIFcalESP) ,5)->
  resYsubjcalESP

#CALIBRACIÓN POR EDAD Y ECIVIL ESPAÑA####
disjunctive(data$P04)->Xecivil
Xedadecivil<-cbind(Xedad,Xecivil)
TedadecivilESP<-c
  (1894440,1980630,1874620,1785945,1991555,4830985,94620,618470)

calib(Xs=Xedadecivil,d=wi,total=TedadecivilESP,method="linear")->
  gpESP2

sum(Ysubj[Ysubj3]*gpESP2[Ysubj3]*wi[Ysubj3])->TYsubjcalESP2
varest(Ys=Ysubj[Ysubj3],Xs=Xedadecivil[Ysubj3,],pik[Ysubj3],w=wi[
  Ysubj3]*gpESP2[Ysubj3])->VTYsubjcalESP2
round(100*(TYsubjcalESP2-1.96*sqrt(VTYsubjcalESP2),TYsubjcalESP2
  +1.96*sqrt(VTYsubjcalESP2))/sum(TedadESP),8)->ICYsubjcalESP2
round(c(TYsubjcalESP2-1.96*sqrt(VTYsubjcalESP2),TYsubjcalESP2+1.96*
  sqrt(VTYsubjcalESP2)),5)->ICYsubjTOTcalESP2
round((TYsubjcalESP2+1.96*sqrt(VTYsubjcalESP2))-(TYsubjcalESP2-1.96*
  sqrt(VTYsubjcalESP2)),5)->ICYsubjTOTDIFcalESP2
round(c(100*TYsubjcalESP2/sum(TedadESP), ICYsubjcalESP2,
  VTYsubjcalESP2, 100*sqrt(VTYsubjcalESP2)/sum(TedadESP), 100*(
  sqrt(VTYsubjcalESP2) /sum(TedadESP)) / (TYsubjcalESP2/sum(
  TedadESP)) , TYsubjcalESP2, ICYsubjTOTcalESP2,
  ICYsubjTOTDIFcalESP2) ,5)-> resYsubjcalESP2

```

Listing A.4: Linear calibration (age, marital status, employment)

```

#CALIBRACIÓN POR EDAD, ESTADO CIVIL Y SITUACIÓN LABORAL ESPAÑA####
disjunctive(data$P05)->Xsitlab
Xedadecivilsitlab<-cbind(Xedad,Xecivil,Xsitlab)
TedadecivilsitlabESP<-c(1894440,1980630,1874620,
1785945,1991555,4830985,94620,618470,4702145,
1836495,165400,708505,52940,70150)

calib(Xs=Xedadecivilsitlab,d=wi,total=TedadecivilsitlabESP,method="
linear")->gpESP3

sum(Ysubj[Ysubj3]*gpESP3[Ysubj3]*wi[Ysubj3])->TYsubjcalESP3
varest(Ys=Ysubj[Ysubj3],Xs=Xedadecivilsitlab[Ysubj3,],pik[Ysubj3
],w=wi[Ysubj3]*gpESP3[Ysubj3])->VTYsubjcalESP3
round(100*(TYsubjcalESP3-1.96*sqrt(VTYsubjcalESP3),TYsubjcalESP3
+1.96*sqrt(VTYsubjcalESP3))/sum(TedadESP),8)->ICYsubjcalESP3
round(c(TYsubjcalESP3-1.96*sqrt(VTYsubjcalESP3),TYsubjcalESP3+1.96*
sqrt(VTYsubjcalESP3)),5)->ICYsubjTOTcalESP3
round((TYsubjcalESP3+1.96*sqrt(VTYsubjcalESP3))-(TYsubjcalESP3-1.96*
sqrt(VTYsubjcalESP3)),5)->ICYsubjTOTDIFcalESP3
round(c(100*TYsubjcalESP3/sum(TedadESP), ICYsubjcalESP3,
VTYsubjcalESP3, 100*sqrt(VTYsubjcalESP3) / sum(TedadESP), 100*(
sqrt(VTYsubjcalESP3) /sum(TedadESP)) / (TYsubjcalESP3 /sum(
TedadESP)), TYsubjcalESP3, ICYsubjTOTcalESP3,
ICYsubjTOTDIFcalESP3),5) ->resYsubjcalESP3

```

Listing A.5: Linear calibration (age, marital status, employment, educational level)

```

#CALIBRACIÓN POR EDAD, ESTADO CIVIL, SITUACIÓN LABORAL Y NIVEL DE
ESTUDIOS ESPAÑA####
disjunctive(data$ESTUDIOS)->Xestudios
Xedadecivilsitlabestudios<-cbind(Xedad,Xecivil,Xsitlab,Xestudios)
TedadecivilsitlabestudiosESP <- c
  (1894440,1980630,1874620,1785945,1991555,4830985,94620,
618470,4702145,1836495,165400,708505,52940,
70150,2760685,2437680,2337270)

calib(Xs=Xedadecivilsitlabestudios,d=wi,total=
  TedadecivilsitlabestudiosESP,method="linear")->gpESP4

sum(Ysubj[Ysubj3]*gpESP4[Ysubj3]*wi[Ysubj3])->TYsubjcalESP4
varest(Ys=Ysubj[Ysubj3],Xs=Xedadecivilsitlabestudios[Ysubj3,],pik[
  Ysubj3],w=wi[Ysubj3]*gpESP4[Ysubj3])->VTYsubjcalESP4
round(100*(TYsubjcalESP4-1.96*sqrt(VTYsubjcalESP4),TYsubjcalESP4
+1.96*sqrt(VTYsubjcalESP4))/sum(TedadESP),8)->ICYsubjcalESP4
round(c(TYsubjcalESP4-1.96*sqrt(VTYsubjcalESP4),TYsubjcalESP4+1.96*
sqrt(VTYsubjcalESP4)),5)->ICYsubjT0TcalESP4
round((TYsubjcalESP4+1.96*sqrt(VTYsubjcalESP4))-(TYsubjcalESP4-1.96*
sqrt(VTYsubjcalESP4)),5)->ICYsubjT0TDIFcalESP4
round(c(100*TYsubjcalESP4/sum(TedadESP), ICYsubjcalESP4,
VTYsubjcalESP4, 100* sqrt(VTYsubjcalESP4) /sum(TedadESP), 100*(
sqrt(VTYsubjcalESP4) /sum(TedadESP))/(TYsubjcalESP4 /sum(
TedadESP)), TYsubjcalESP4, ICYsubjT0TcalESP4,
ICYsubjT0TDIFcalESP4),5)-> resYsubjcalESP4

```

Listing A.6: Linear calibration (age, marital status, employment, educational level, country of birth)

```

#CALIBRACIÓN POR EDAD, ESTADO CIVIL, SITUACIÓN LABORAL, NIVEL DE
ESTUDIOS y PAÍS NACIMIENTO ESPAÑA####

disjunctive(data$P02_1BIS) -> Xnac
Xedadecivilsitlabestudiosnac <- cbind(Xedad, Xecivil, Xsitlab, Xestudios,
Xnac)
TedadecivilsitlabestudiosnacESP <- c
(1894440, 1980630, 1874620, 1785945, 1991555, 4830985, 94620,
618470, 4702145, 1836495, 165400, 708505, 52940, 70150,
2760685, 2437680, 2337270, 6238395, 416120, 881115)

calib(Xs=Xedadecivilsitlabestudiosnac, d=wi, total=
TedadecivilsitlabestudiosnacESP, method="linear") -> gpESP5

sum(Ysubj[Ysubj3]*gpESP5[Ysubj3]*wi[Ysubj3]) -> TYsubjcaESP5
varest(ys=Ysubj[Ysubj3], Xs=Xedadecivilsitlabestudiosnac[Ysubj3, ],
pik[Ysubj3], w=wi[Ysubj3]*gpESP5[Ysubj3]) -> VTYsubjcaESP5
round(100*(TYsubjcaESP5-1.96*sqrt(VTYsubjcaESP5), TYsubjcaESP5
+1.96*sqrt(VTYsubjcaESP5))/sum(TedadESP), 8) -> ICYsubjcaESP5
round(c(TYsubjcaESP5-1.96*sqrt(VTYsubjcaESP5), TYsubjcaESP5+1.96*
sqrt(VTYsubjcaESP5)), 5) -> ICYsubjTOTcaESP5
round((TYsubjcaESP5+1.96*sqrt(VTYsubjcaESP5)) - (TYsubjcaESP5-1.96*
sqrt(VTYsubjcaESP5)), 5) -> ICYsubjTOTDIFcaESP5
round(c(100*TYsubjcaESP5/sum(TedadESP), ICYsubjcaESP5,
VTYsubjcaESP5, 100*sqrt(VTYsubjcaESP5)/sum(TedadESP), 100*(
sqrt(VTYsubjcaESP5)/sum(TedadESP)) / (TYsubjcaESP5/sum(TedadESP
)), TYsubjcaESP5, ICYsubjTOTcaESP5, ICYsubjTOTDIFcaESP5), 5) ->
resYsubjcaESP5

```

Listing A.7: Linear calibration (age, marital status, employment, educational level, country of birth, number of children)

```
#CALIBRACIÓN POR EDAD, ESTADO CIVIL, SITUACIÓN LABORAL, NIVEL DE
ESTUDIOS, PAÍS NACIMIENTO y NÚMERO HIJOS ESPAÑA####

disjunctive(data$p30bis)->Xhijos
Xedadecivilsitlabestudiosnachijos<-cbind(Xedad,Xecivil,Xsitlab,
Xestudios,Xnac,Xhijos)
TedadecivilsitlabestudiosnachijosESP <- c
(1894440,1980630,1874620,1785945,1991555,4830985,94620,
618470,4702145,1836495,165400,708505,52940,70150,
2760685,2437680,2337270,6238395,416120,881115,
2262245,1908390,2663355,539945,161695)

calib(Xs=Xedadecivilsitlabestudiosnachijos,d=wi,total=
TedadecivilsitlabestudiosnachijosESP,method="linear")->gpESP6

sum(Ysubj[Ysubj3]*gpESP6[Ysubj3]*wi[Ysubj3])->TYsubjcalESP6
varest(Ys=Ysubj[Ysubj3],Xs=Xedadecivilsitlabestudiosnachijos[
Ysubj3,],pik[Ysubj3],w=wi[Ysubj3]*gpESP6[Ysubj3])->
VTYsubjcalESP6
round(100*(TYsubjcalESP6-1.96*sqrt(VTYsubjcalESP6),TYsubjcalESP6
+1.96*sqrt(VTYsubjcalESP6))/sum(TedadESP),8)->ICYsubjcalESP6
round(c(TYsubjcalESP6-1.96*sqrt(VTYsubjcalESP6),TYsubjcalESP6+1.96*
sqrt(VTYsubjcalESP6)),5)->ICYsubjT0TcalESP6
round((TYsubjcalESP6+1.96*sqrt(VTYsubjcalESP6))-(TYsubjcalESP6-1.96*
sqrt(VTYsubjcalESP6)),5)->ICYsubjT0TDIFcalESP6
round(c(100*TYsubjcalESP6/sum(TedadESP), ICYsubjcalESP6,
VTYsubjcalESP6, 100*sqrt(VTYsubjcalESP6) /sum(TedadESP), 100*(
sqrt(VTYsubjcalESP6) /sum(TedadESP))/ (TYsubjcalESP6/sum(
TedadESP)), TYsubjcalESP6, ICYsubjT0TcalESP6,
ICYsubjT0TDIFcalESP6),5) ->resYsubjcalESP6
```

Listing A.8: Calibration estimations for infertility

```

# ESP RESULTADOS: TODAS LAS PREVALENCIAS SIN CALIBRAR (SUPOB Y POB)
  Y CALIBRADAS####
cbind(resYsubj, resYsubjPOB, resYsubjcalESP, resYsubjcalESP2,
      resYsubjcalESP3, resYsubjcalESP4, resYsubjcalESP5, resYsubjcalESP6)
->reswisubjcalESP
reswisubjcalESP

### ESTIMADORES: [1] PROPORCIÓN, [2] MIN IC PROPORCIÓN, [3] MAX IC
  PROPORCIÓN, [4] VARIANZA, [5] ERROR MUESTRAL, [6] CV, [7] TOTAL,
  [8] MIN IC TOTAL, [9] MAX IC TOTAL, [10] LONGITUD TOTAL.
# Vectores auxiliares: edad; ESP2:edad y ecivil; ESP3: edad, ecivil
  y sitlab; ESP4: edad, ecivil, sitlab y estudios; ESP5: edad,
  ecivil, sitlab, estudios y nacimiento; ESP6: edad, ecivil,
  sitlab, estudios, nacimiento y hijos

resYsubj  resYsubjPOB  resYsubjcalESP  resYsubjcalESP2  resYsubjcalESP3
resYsubjcalESP4  resYsubjcalESP5  resYsubjcalESP6
[1,]      9.50267      9.41315      9.38356      9.12050
      8.98329      8.87379      8.38969      8.22076
[2,]      6.46858      6.40763      9.37371      9.11085
      8.97370      8.86426      8.38009      8.21129
[3,]     12.53677     12.41866      9.39342      9.13016
      8.99287      8.88331      8.39929      8.23023
[4,] 146441.08000 146441.08000 143624.45968 137855.20733
      135826.49542 134058.99359 136205.22638 132661.84819
[5,]      1.54801      1.53343      0.00503      0.00493
      0.00489      0.00486      0.00490      0.00483
[6,]     16.29025     16.29025      0.05360      0.05402
      0.05444      0.05475      0.05838      0.05880
[7,]   2349.11178   2349.11178   707111.08476   687287.83695
      676947.84115   668696.06803   632216.41794   619486.51430
[8,]   1599.06641   1599.06641   706368.28753   686560.11134
      676225.49008   667978.43230   631493.06049   618772.62792
[9,]   3099.15715   3099.15715   707853.88199   688015.56256
      677670.19221   669413.70375   632939.77539   620200.40067
[10,]  1500.09073   1500.09073   1485.59446   1455.45122
      1444.70214   1435.27144   1446.71490   1427.77275

```

Listing A.9: Percentage reductions in variance and prevalence estimations after calibration

```

###ESP VARIACIÓN PORCENTUAL DE LA VARIANZA###
# ESP: edad; ESP2:edad y ecivil; ESP3: edad, ecivil y sitlab; ESP4:
  edad, ecivil, sitlab y estudios; ESP5: edad, ecivil, sitlab,
  estudios y nacimiento; ESP6: edad, ecivil, sitlab, estudios,
  nacimiento y hijos

round(((VTYsubjcalESP/VTYsubj)-1)*100,2)->VTYsubjvariacionESP
round(((VTYsubjcalESP2/VTYsubj)-1)*100,2)->VTYsubjvariacionESP2
round(((VTYsubjcalESP3/VTYsubj)-1)*100,2)->VTYsubjvariacionESP3
round(((VTYsubjcalESP4/VTYsubj)-1)*100,2)->VTYsubjvariacionESP4
round(((VTYsubjcalESP5/VTYsubj)-1)*100,2)->VTYsubjvariacionESP5
  round(((VTYsubjcalESP6/VTYsubj)-1)*100,2)->VTYsubjvariacionESP6
cbind(VTYsubjvariacionESP,VTYsubjvariacionESP2,VTYsubjvariacionESP3,
  VTYsubjvariacionESP4,VTYsubjvariacionESP5,VTYsubjvariacionESP6)
-> VTYsubjvariacionESPglobal

VTYsubjvariacionESP VTYsubjvariacionESP2 VTYsubjvariacionESP3
  VTYsubjvariacionESP4 VTYsubjvariacionESP5 VTYsubjvariacionESP6
-1.92 -5.86 -7.25 -8.46 -6.99 -9.41

###VARIACIÓN PORCENTUAL DE LA PROPORCIÓN###

round((TYsubjcalESP/sum(TedadESP)/(TYsubj/sum(wi))-1)*100,2)->
  PYsubjvariacionESP
round((TYsubjcalESP2/sum(TedadESP)/(TYsubj/sum(wi))-1)*100,2)->
  PYsubjvariacionESP2
round((TYsubjcalESP3/sum(TedadESP)/(TYsubj/sum(wi))-1)*100,2)->
  PYsubjvariacionESP3
round((TYsubjcalESP4/sum(TedadESP)/(TYsubj/sum(wi))-1)*100,2)->
  PYsubjvariacionESP4
round((TYsubjcalESP5/sum(TedadESP)/(TYsubj/sum(wi))-1)*100,2)->
  PYsubjvariacionESP5
round((TYsubjcalESP6/sum(TedadESP)/(TYsubj/sum(wi))-1)*100,2)->
  PYsubjvariacionESP6
cbind(PYsubjvariacionESP,PYsubjvariacionESP2,PYsubjvariacionESP3,
  PYsubjvariacionESP4,PYsubjvariacionESP5,PYsubjvariacionESP6) ->
  PYsubjvariacionESPglobal

PYsubjvariacionESP PYsubjvariacionESP2 PYsubjvariacionESP3
  PYsubjvariacionESP4 PYsubjvariacionESP5 PYsubjvariacionESP6
-0.31 -3.11 -4.57 -5.73 -10.87 -12.67

```

Listing A.10: Disabling Chronic Pain: starting R

```

### PREPARACIÓN R ###
setwd("~/Dropbox/DCD CALIBRADO DROPBOX")
library(foreign)
data<-read.spss("data.sav",to.data.frame=T, use.value.labels=F)
library("sampling",lib.loc="/Library/Frameworks/R.framework/Versions
/3.2/Resources/library")
library("samplingVarEst",lib.loc="/Library/Frameworks/R.framework/
Versions/3.2/Resources/library")

###Definiendo las variables
data$uno <- 1
data$uno->uno

#Definiendo los pesos
data$elevacion->wi
pik<-1/wi

#Para evitar la notación científica e imprimir con 7 decimales
options(scipen=100,digits=8)

```

Listing A.11: Estimation based on sampling design weights

```

####ESTIMADOR SIN CALIBRAR PARA DOLOR CRÓNICO DISCAPACITANTE####
#Diseño muestral CR (con reposición) y corrección para poblaciones
finitas bajo MAS;
#Para evitar la notación científica e imprimir con 7 decimales
options(scipen=100,digits=8)
sum(Y*wi)->TY
varest(Ys=Y,Xs=NULL,pik,w=wi)->VTY
100*c(TY-1.96*sqrt(VTY),TY+1.96*sqrt(VTY))/sum(wi)->ICY c(TY-1.96*
sqrt(VTY),TY+1.96*sqrt(VTY))->ICYTOT (TY+1.96*sqrt(VTY))-(TY
-1.96*sqrt(VTY))->ICYTOTDIF
c(100*TY/sum(wi), ICY, 100*sqrt(VTY)/sum(wi), 100*(sqrt(VTY)/sum(wi)
)/(TY/sum(wi)), TY, ICYTOT, ICYTOTDIF, VTY)->resYHT

### ESTIMADORES RESULTADOS: #[1]PREVALENCIA (%), MIN IC PREVALENCIA,
MAX IC PREVALENCIA, ERROR MUESTRAL DE LA MEDIA,CV(%), TOTAL,
MIN IC TOTAL, [8]MAX IC TOTAL, LONGITUD TOTAL,VARIANZA DEL TOTAL
.

#VALORES DE resYHT
#[1] 11.50717199 10.73070007 12.28364392
0.39615915 3.44271509 793455.32317615
739915.16712297
#[8] 846995.47922934 107080.31210636 746186044.92892361

```

Listing A.12: Linear calibration (marginals)

```

#CALIBRACIÓN LINEAL MARGINAL DCD POR EDADG, SEXO Y ESTUDIOS ANDALUCÍA
####
disjunctive(data$EDADG) ->Xedad
disjunctive(data$SEX0) ->Xsexo
disjunctive(data$ESTUDIOS) ->Xestudios
Xedadsexoestudios<- cbind(Xedad,Xsexo,Xestudios)
Tedad<-c(3536345,2072575,1260880) Tedadsexoestudios<-c
(3536345,2072575,1260880,3378625,3491175,2101280,3618590,
1149930)

calib(Xs=Xedadsexoestudios,d=wi,total=Tedadsexoestudios,method="
linear") ->gp2

sum(Y*gp2*wi) ->TYcal
varest(Ys=Y,Xs=Xedadsexoestudios,pik,w=wi*gp2) ->VTYcal 100*c(TYcal
-1.96*sqrt(VTYcal),TYcal+1.96*sqrt(VTYcal))/sum(Tedad) ->ICYcal c
(TYcal-1.96*sqrt(VTYcal),TYcal+1.96*sqrt(VTYcal)) ->ICYT0Tcal (
TYcal+1.96*sqrt(VTYcal)) - (TYcal-1.96*sqrt(VTYcal)) ->ICYT0TDIFcal
c(100*TYcal/sum(Tedad), ICYcal, 100*sqrt(VTYcal)/sum(Tedad), 100*(
sqrt(VTYcal)/sum(Tedad))/(TYcal/sum(Tedad)), TYcal, ICYT0Tcal,
ICYT0TDIFcal, VTYcal) ->resYcalineal

#VALORES resYcalineal
#[1] 10.75857832 10.01213190 11.50502473
0.38084001 3.53987298 739092.81327556
687813.43755890
#[8] 790372.18899222 102558.75143332 684499784.95682335

```

Listing A.13: Linear calibration (crossings)

```

#CALIBRACIÓN LINEAL CRUCES DCD POR EDAD, SEXO Y ESTUDIOS####
data$edadsexoestudios <- factor(data$EDADG):factor(data$SEXO):
  factor(data$ESTUDIOS)
Tedadsexoestudioscruce <- c
  (307360,1192015,305110,217325,1094065,420470,316005,
  540110,172950,359070,521950,162490,351850,138035,
  55190,549670,132415,33720)
disjunctive(data$edadsexoestudios)->Xedadsexoestudioscruce

calib(Xs=Xedadsexoestudioscruce,d=wi,total=Tedadsexoestudioscruce,
  method="raking")->gp2linealcruces

sum(Y*gp2linealcruces*wi)->TYcal varest(Ys=Y,Xs=Xedadsexoestudios,
  pik,w=wi*gp2linealcruces)->VTYcal 100*c(TYcal-1.96*sqrt(VTYcal),
  TYcal+1.96*sqrt(VTYcal))/sum(Tedad)->ICYcal c(TYcal-1.96*sqrt(
  VTYcal),TYcal+1.96*sqrt(VTYcal))->ICYT0Tcal (TYcal+1.96*sqrt(
  VTYcal))-(TYcal-1.96*sqrt(VTYcal))->ICYT0TDIFcal
c(100*TYcal/sum(Tedad), ICYcal, 100*sqrt(VTYcal)/sum(Tedad), 100*(
  sqrt(VTYcal)/sum(Tedad))/(TYcal/sum(Tedad)), TYcal, ICYT0Tcal,
  ICYT0TDIFcal, VTYcal)->resYcallinealcruce

#VALORES DE resYcallinealcruce
#[1]      10.90751130      10.16096672      11.65405588
      0.38089009      3.49199813      749324.21115220
      698038.09169973
#[8]      800610.33060467      102572.23890494      684679833.53112984

```

Listing A.14: Raking calibration (marginals and crossings)

```

#CALIBRACIÓN RAKING MARGINAL DCD POR EDAD, SEXO Y ESTUDIOS####
calib(Xs=Xedadsexoestudios,d=wi,total=Tedadsexoestudios,method="
  linear")->gp2raking

sum(Y*gp2raking*wi)->TYcal
varest(Ys=Y,Xs=Xedadsexoestudios,pik,w=wi*gp2raking)->VTYcal 100*c(
  TYcal-1.96*sqrt(VTYcal),TYcal+1.96*sqrt(VTYcal))/sum(Tedad)->
  ICYcal c(TYcal-1.96*sqrt(VTYcal),TYcal+1.96*sqrt(VTYcal))->
  ICYT0Tcal (TYcal+1.96*sqrt(VTYcal))-(TYcal-1.96*sqrt(VTYcal))->
  ICYT0TDIFcal

c(100*TYcal/sum(Tedad), ICYcal, 100*sqrt(VTYcal)/sum(Tedad), 100*(
  sqrt(VTYcal)/sum(Tedad))/(TYcal/sum(Tedad)), TYcal, ICYT0Tcal,
  ICYT0TDIFcal, VTYcal)->resYcalrakingmarginales

#VALORES DE resYcalrakingmarginales
#[1]      10.75857832      10.01213190      11.50502473
        0.38084001      3.53987298      739092.81327556
        687813.43755890
#[8]      790372.18899222      102558.75143332 684499784.95682335

#CALIBRACIÓN RAKING CRUCES DCD POR EDAD, SEXO Y ESTUDIOS#### calib(
  Xs=Xedadsexoestudioscruce,d=wi,total=Tedadsexoestudioscruce,
  method="raking")->gp2rakingcruces

sum(Y*gp2rakingcruces*wi)->TYcal varest(Ys=Y,Xs=Xedadsexoestudios,
  pik,w=wi*gp2rakingcruces)->VTYcal 100*c(TYcal-1.96*sqrt(VTYcal),
  TYcal+1.96*sqrt(VTYcal))/sum(Tedad)->ICYcal
c(TYcal-1.96*sqrt(VTYcal),TYcal+1.96*sqrt(VTYcal))->ICYT0Tcal (TYcal
  +1.96*sqrt(VTYcal))-(TYcal-1.96*sqrt(VTYcal))->ICYT0TDIFcal

c(100*TYcal/sum(Tedad), ICYcal, 100*sqrt(VTYcal)/sum(Tedad), 100*(
  sqrt(VTYcal)/sum(Tedad))/(TYcal/sum(Tedad)), TYcal, ICYT0Tcal,
  ICYT0TDIFcal, VTYcal)->resYcalrakingcruce

#VALORES DE resYcalrakingcruce
#[1]      10.90751130      10.16096672      11.65405588
        0.38089009      3.49199813      749324.21115220
        698038.09169973
#[8]      800610.33060467      102572.23890494 684679833.53112984

```

Listing A.15: Model-assisted calibration

```

####CALIBRACIÓN DCD MODELO ASISTIDA LINEAL POR EDAD, SEXO Y ESTUDIOS
      MEDIANTE ÚNICA ECUACIÓN DE CALIBRACIÓN####
###En Xs se introducen las probabilidades de tener dolor crónico
      discapacitante y en total la suma total de esas probabilidades

TOTAL <- as.factor(data$CENSO*data$DCDprob)
#Esta suma es 727368.1 que difiere algo de la de sin calibrar
      (793455.32)

sumTOTAL <- sum(as.numeric(levels(TOTAL)))
TOTALNODCD <- as.factor(data$CENSO*data$nDCDnDCprob)
sumTOTALNODCD <- sum(as.numeric(levels(TOTALNODCD)))

#sumTOTALNODCD+sumTOTAL = 6869800 que coincide con la del CENSO

calib(Xs=data$DCDprob,d=wi,total=sumTOTAL, method="linear",
      description=FALSE)->gp2loglineal

sum(Y*gp2loglineal*wi)->TYcal varest(Ys=Y,Xs=Xedadsexoestudios,pik,w
      =wi*gp2loglineal)->VTYcal 100*c(TYcal-1.96*sqrt(VTYcal),TYcal
      +1.96*sqrt(VTYcal))/sum(Tedad)->ICYcal c(TYcal-1.96*sqrt(VTYcal)
      ,TYcal+1.96*sqrt(VTYcal))->ICYT0Tcal (TYcal+1.96*sqrt(VTYcal))-
      (TYcal-1.96*sqrt(VTYcal))->ICYT0TDIFcal
c(100*TYcal/sum(Tedad), ICYcal, 100*sqrt(VTYcal)/sum(Tedad), 100*(
      sqrt(VTYcal)/sum(Tedad))/(TYcal/sum(Tedad)), TYcal, ICYT0Tcal,
      ICYT0TDIFcal, VTYcal)->resYcalloglineal

#VALORES DE resYcalloglineal
# [1]      10.77564915      10.02924098      11.52205732
      0.38082049      3.53408401      740265.54535536
      688988.79708913
#[8]      791542.29362160      102553.49653247 684429642.01339519

```

Listing A.16: Model-assisted calibration (raking)

```

#####CALIBRACIÓN DCD MODELO ASISTIDA RAKING POR EDAD, SEXO Y ESTUDIOS
MEDIANTE ÚNICA ECUACIÓN DE CALIBRACIÓN#####

calib(Xs=data$DCDprob,d=wi,total=sumTOTAL, method="raking",
description=FALSE) ->gp2lograking

sum(Y*gp2lograking*wi) ->TYcal varest(Ys=Y,Xs=Xedadsexoestudios,pik,w
=wi*gp2lograking) ->VTYcal 100*c(TYcal-1.96*sqrt(VTYcal),TYcal
+1.96*sqrt(VTYcal))/sum(Tedad) ->ICYcal c(TYcal-1.96*sqrt(VTYcal)
,TYcal+1.96*sqrt(VTYcal)) ->ICYT0Tcal (TYcal+1.96*sqrt(VTYcal))-(
TYcal-1.96*sqrt(VTYcal)) ->ICYT0TDIFcal
c(100*TYcal/sum(Tedad), ICYcal, 100*sqrt(VTYcal)/sum(Tedad), 100*(
sqrt(VTYcal)/sum(Tedad))/(TYcal/sum(Tedad)), TYcal, ICYT0Tcal,
ICYT0TDIFcal, VTYcal) ->resYcallograking

#VALORES DE resYcal raking modelo-asistida
#[1] 10.77517247 10.02876443 11.52158050
0.38082043 3.53423974 740232.79809758
688956.05871136
#[8] 791509.53748380 102553.47877244 684429404.95683587

```

Listing A.17: Calibration estimations for disabling chronic pain

```

#RESULTADOS: TODAS LAS PREVALENCIAS SIN CALIBRAR Y CALIBRADAS####
round(cbind(resYHT,resYcalineal,resYcallinealcruce,resYcalloglineal)
,8)->reswiYcal write.csv (reswiYcal,file="reswiYcal.csv", row.
names=FALSE,fileEncoding = "latin1")

### ESTIMADORES: [1] PROPORCIÓN, [2] MIN IC PROPORCIÓN, [3] MAX IC
PROPORCIÓN, [4] ERROR MUESTRAL, [5] CV, [6] TOTAL, [7] MIN IC
TOTAL, [8] MAX IC TOTAL, [9] LONGITUD TOTAL, [10] VARIANZA. #
      resYHT      resYcalineal resYcallinealcruce
resYcalloglineal #[1,]      11.50717199      10.75857832
      10.90751130      10.77564915
#[2,]      10.73070007      10.01213190      10.16096672
      10.02924098
#[3,]      12.28364392      11.50502473      11.65405588
      11.52205732
#[4,]      0.39615915      0.38084001      0.38089009
      0.38082049
#[5,]      3.44271509      3.53987298      3.49199813
      3.53408401
#[6,]      793455.32317615      739092.81327556      749324.21115220
      740265.54535536
#[7,]      739915.16712297      687813.43755890      698038.09169973
      688988.79708913
#[8,]      846995.47922934      790372.18899222      800610.33060467
      791542.29362160
#[9,]      107080.31210636      102558.75143332      102572.23890494
      102553.49653247
#[10,] 746186044.92892361 684499784.95682335 684679833.53112984
      684429642.01339519

```

Listing A.18: Percentage reductions in variance and prevalence estimations after calibration

```

### VARIACIÓN PORCENTUAL DE LA VARIANZA###
((resYcalineal[10]/resYHT[10])-1)*100->VTYvariacioncalineal ((
  resYcallinealcruce[10]/resYHT[10])-1)*100->
  VTYvariacioncallinealcruce ((resYcalrakingmarginales[10]/resYHT
[10])-1)*100->VTYvariacioncalrakingmarginales ((resYcalloglineal
[10]/resYHT[10])-1)*100->VTYvariacioncalloglineal ((
  resYcallograking[10]/resYHT[10])-1)*100->
  VTYvariacioncallograking
cbind(VTYvariacioncalineal,VTYvariacioncallinealcruce,
  VTYvariacioncalloglineal)->VTYvariacion
#-8.266874          -8.2427448          -8.2762742

###VARIACIÓN PORCENTUAL DE LA PREVALENCIA###
((resYcalineal[1]/resYHT[1])-1)*100->PYvariacioncalineal ((
  resYcallinealcruce[1]/resYHT[1])-1)*100->
  PYvariacioncallinealcruce ((resYcalrakingmarginales[1]/resYHT
[1])-1)*100->PYvariacioncalrakingmarginales ((resYcalloglineal
[1]/resYHT[1])-1)*100->PYvariacioncalloglineal ((
  resYcallograking[1]/resYHT[1])-1)*100->PYvariacioncallograking
cbind(PYvariacioncalineal,PYvariacioncallinealcruce,
  PYvariacioncalloglineal)->PYvariacion #-6.5054531
          -5.2111909          -6.3571036

###VARIACIÓN PORCENTUAL DEL TOTAL POBLACIONAL###
(resYcalineal[7]-resYHT[7])->TYvariacioncalineal (resYcallinealcruce
[7]-resYHT[7])->TYvariacioncallinealcruce (
  resYcalrakingmarginales[7]-resYHT[7])->
  TYvariacioncalrakingmarginales (resYcalloglineal[7]-resYHT[7])->
  TYvariacioncalloglineal (resYcallograking[7]-resYHT[7])->
  TYvariacioncallograking
cbind(TYvariacioncalineal,TYvariacioncallinealcruce,
  TYvariacioncalloglineal)->TYvariacion #-52101.73
          -41877.075          -50926.37

variacion <- data.frame(VTYvariacion,PYvariacion,TYvariacion)
write.csv (variacion,file="variacion.csv", row.names=FALSE,
  fileEncoding = "latin1")

```


BIBLIOGRAFÍA

- Aitchison, J y CGG Aitken
1976 "Multivariate binary discrimination by the kernel method", *Biometrika*, 63, pág. 413-420. (Citado en la pág. 49.)
- Alfons, A, J Holzer y M Templ
2015 *R package laeken: Estimation of indicators on social exclusion and poverty*, pág. 71, <https://cran.r-project.org/web/packages/laeken/index.html>. (Citado en la pág. 93.)
- Altman, Russ, Philip Pizzo, Robert Gibbons, Kathy Hudson, Renee Jenkins y Brendan Lee
2014 *National Children's Study (NCS) Working Group. Final Report*. Government Document, http://acd.od.nih.gov/reports/NCS_WG_FINAL_REPORT.pdf. (Citado en la pág. 3.)
- Álvarez-González, J, A AyusoFernández y F Caba-Barrientos
2010 *Plan Andaluz de atención a las personas con dolor: 2010–2013*, Consejería de Salud, Junta de Andalucía. Sevilla, http://www.juntadeandalucia.es/export/drupaljda/plan_atencion_dolor-1_0.pdf. (Citado en la pág. 140.)
2014 *IASP Taxonomy: Pain Terms*, International Association for the Study of Pain (IASP). Washington, <http://www.iasp-pain.org/Taxonomy?navItemNumber=576#Pain>. (Citado en las pág. 140, 141.)
- Andersson, Claes y Lennart Nordberg
2000 *CLAN - A SAS-Program for Computation of Point-and Standard Error Estimates in Sample Surveys*, American Statistical Association, Buffalo, New York, <https://www.amstat.org/meetings/ices/2000/proceedings/S09.pdf>. (Citado en la pág. 91.)
- Arcos, A, JM Contreras y MM Rueda
2014 "A Novel Calibration Estimator in Social Surveys", *Sociological Methods & Research*, 43, 3, pág. 465-489, DOI: [10.1177/0049124113507906](https://doi.org/10.1177/0049124113507906). (Citado en las pág. 107, 119, 122.)
- Arcos, Antonio, D Molina, M Giovanna Ranalli y María del Mar Rueda
2015 "Frames2: A Package for Estimation in Dual Frame Surveys", *The R Journal*, 7/1, pág. 52-72, <https://journal.r-project.org/archive/2015-1/arcos-molina-ranalli-et-al.pdf>. (Citado en las pág. 75, 92.)

- Arcos, Antonio, María del Mar Rueda y D Molina
 2013 "Calibration in Complex Survey using R", *The R User Conference, useR! 2013 July 10-12 2013 University of Castilla-La Mancha, Albacete, Spain*, 10, 30, pág. 140. (Citado en la pág. 98.)
- Arcos, Antonio, María del Mar Rueda, M Giovanna Ranalli y D Molina
 2015 *R package Frames2: Estimation in Dual Frame Surveys*, pág. 60, <https://cran.r-project.org/web/packages/Frames2/index.html>. (Citado en las pág. 92, 93.)
- Arora, V. S., M. Karanikolos, A. Clair, A. Reeves, D. Stuckler y M. McKee
 2015 "Data Resource Profile: The European Union Statistics on Income and Living Conditions (EU-SILC)", *International Journal of Epidemiology*, 44, 2, pág. 451-461, DOI: [10.1093/ije/dyv069](https://doi.org/10.1093/ije/dyv069), [%3CGo%20to%20ISI%3E://WOS:000357106100010](https://pubmed.ncbi.nlm.nih.gov/26003571/). (Citado en la pág. 4.)
- Asmundson, GJ y J Katz
 2009 "Understanding the co-occurrence of anxiety disorders and chronic pain: state-of-the-art", *Depress Anxiety*, 26, 10, pág. 888-901. (Citado en la pág. 142.)
- Australian Health Technology Advisory Comitee
 1998 *Review of assisted reproductive technology*, Generic. (Citado en la pág. 103.)
- Ayala, A, C Rodríguez-Blázquez, B Frades-Payo, MJ Forjaza, P Martínez-Martín, G Fernández-Mayoralas y F Rojo-Pérez
 2012 "Psychometric properties of the Functional Social Support Questionnaire and the Loneliness Scale in non-institutionalized older adults in Spain", *GacSanit*, 26, 4, pág. 317-24. (Citado en la pág. 124.)
- Azevedo, LF, A Costa-Pereira, L Mendonça, CC Dias y JM Castro-Lopes
 2012 "Epidemiology of chronic pain: a population-based nationwide study on its prevalence and characteristics and associated disability in Portugal", *J Pain*, 13, 8, pág. 773-83. (Citado en las pág. 121, 122, 140, 142.)
- Azorín, F. y J.L Sánchez-Crespo
 1995 *Métodos y aplicaciones del muestreo*, Alianza Editorial, pág. 400. (Citado en la pág. 12.)
- Bair, MJ, J Wu, TM Damush, JM Sutherland y K Kroenke
 2008 "Association of depression and anxiety alone and in combination with chronic musculoskeletal pain in primary care patients", *Psychosom Med*, 70, 8, pág. 890-97. (Citado en la pág. 142.)

Bankier, MD

- 1986 "Estimators based on several stratified samples with applications to multiple frame surveys", *Journal of the American Statistical Association*, 81, pág. 1074-1079. (Citado en la pág. 86.)

Barranco-Chamorro, I, MD Jiménez-Gamero, JA Mayor Gallego y JL Moreno-Rebollo

- 2015 "Case-deletion diagnostic for penalized calibration estimators and BLUP under linear mixed models in survey sampling", *Computational Statistics and Data Analysis*, 87, pág. 18-33. (Citado en la pág. 75.)

Barranco-Chamorro, I, MD Jiménez-Gamero, JL Moreno-Rebollo y JM Muñoz-Pichardo

- 2012 "Case-deletion type diagnostics for calibration estimators in survey sampling", *Computational Statistics and Data Analysis*, 56, 7, pág. 2219-2236. (Citado en la pág. 75.)

Baune, B, R Caniato, MA Garcia-Alcaraz y K Berger

- 2008 "Combined effects of major depression, pain and somatic disorders on general functioning in the general adult population", *Pain*, 138, 2, pág. 310-7. (Citado en la pág. 142.)

Bedoya, JM

- 1994 "Estimación de la prevalencia de esterilidad/infertilidad a partir del censo de población en España en 1991", *Prog Obst Gin*, 37, pág. 557-562. (Citado en la pág. 103.)

Bhattacharya, S., M. Porter, E. Amalraj, A. Templeton, M. Hamilton, A. J. Lee y J. J. Kurinczuk

- 2009 "The epidemiology of infertility in the North East of Scotland", en *Hum Reprod*, England, vol. 24, pág. 3096-107, DOI: [10.1093/humrep/dep287](https://doi.org/10.1093/humrep/dep287). (Citado en las pág. 103, 116, 118.)

Biemer, P.P. y L.E. Lyberg

- 2003 *Introduction to survey quality*, John Wiley & Sons, New Jersey. (Citado en la pág. 14.)

Bingefors, K y D Isacson

- 2004 "Epidemiology and co-morbidity and and impact on health-related quality of life of self-reported headache and musculoskeletal pain – a gender perspective", *Eur J Pain*, 8, pág. 435-50. (Citado en la pág. 140.)

- Björnsdóttir, SV, SH Jónsson y Valdimarsdóttir UA
 2013 "Functional limitations and physical symptoms of individuals with chronic pain", *Scand J Rheumatol*, 42, pág. 59-70. (Citado en la pág. 141.)
- Blyth, FM, LM March, AJ Brnabic, LR Jorm, M Williamson y MJ Cousins
 2006 "Chronic pain in Australia: a prevalence study", *Pain*, 89, 2, pág. 127-34. (Citado en las pág. 121, 122, 140-142.)
- Blyth, FM, LM March, AJM Brnabic y Cousins MJ
 2014 "Chronic pain and frequent use of health care", *Pain*, 111, 1, pág. 51-8. (Citado en la pág. 122.)
- Boersch-Supan, Axel, Martina Brandt, Christian Hunkler, Thorsten Kneip, Julie Korbmacher, Frederic Malter, Barbara Schaan, Stephanie Stuck, Sabrina Zuber y Share Cent Coordination Team
 2013 "Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE)", *International Journal of Epidemiology*, 42, 4, pág. 992-1001, DOI: [10.1093/ije/dyt088](https://doi.org/10.1093/ije/dyt088), [%3CGo%20to%20ISI%3E://WOS:000325167800011](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3251678/). (Citado en la pág. 4.)
- Boivin, J., L. Bunting, J. A. Collins y K. G. Nygren
 2007 "International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care", en *Hum Reprod*, England, vol. 22, pág. 1506-12, DOI: [10.1093/humrep/dem046](https://doi.org/10.1093/humrep/dem046). (Citado en la pág. 103.)
- Bolumar, F., J. Olsen y J. Boldsen
 1996 "Smoking reduces fecundity: a European multicenter study on infertility and subfecundity. The European Study Group on Infertility and Subfecundity", *Am J Epidemiol*, 143, 6, pág. 578-87. (Citado en la pág. 104.)
- Breidt, J y JD Opsomer
 2000 "Local polynomial regression estimators in survey sampling", *The Annals of Statistics*, 28, pág. 1026-1053. (Citado en las pág. 45, 47, 50.)
- Breivik, H, B Collett, V Ventafridda, R Cohen y D Gallacher
 2006 "Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment", *Eur J Pain*, 10, 4, pág. 287-333. (Citado en las pág. 121, 140, 141, 143.)
- Brewer, KRW
 1999 "Cosmetic Calibration with Unequal Probability Sampling", *Survey Methodology*, 25, pág. 205-212. (Citado en la pág. 64.)

- Bruehl, S, OY Chung, JN Jirjis y S Biridepalli
 2005 "Prevalence of clinical hypertension in patients with chronic pain compared to non pain general medical patients", *Clin J Pain*, 21, 2, pág. 147-53. (Citado en la pág. 142.)
- Buck Louis, G. M., R. Sundaram, E. F. Schisterman, A. M. Sweeney, C. D. Lynch, R. E. Gore-Langton, Z. Chen, S. Kim, K. L. Caldwell y D. B. Barr
 2012 "Heavy metals and couple fecundity, the LIFE Study", *Chemosphere*, 87, 11, pág. 1201-7, DOI: [10.1016/j.chemosphere.2012.01.017](https://doi.org/10.1016/j.chemosphere.2012.01.017), <http://www.ncbi.nlm.nih.gov/pubmed/22309709>. (Citado en la pág. 118.)
- Cabrera-León, A, A Daponte y Subprograma Crisis y Salud CIBERESP
 2016 "Indicadores contextuales para evaluar los determinantes sociales de salud y la crisis económica española", *Gaceta Sanitaria, en revisión*. (Citado en la pág. 69.)
- Cabrera-León, A., V. Lopez-Villaverde, M. Rueda y M. N. Moya-Garrido
 2015 "Calibrated prevalence of infertility in 30- to 49-year-old women according to different approaches: a cross-sectional population-based study", *Human Reproduction (FI=4.6; D1; Ranking: 5/79)*, 30, 11, pág. 2677-2685, DOI: [10.1093/humrep/dev226](https://doi.org/10.1093/humrep/dev226), <http://www.ncbi.nlm.nih.gov/pubmed/26370663>. (Citado en las pág. VIII, XII, 74, 103, 122, 143, 145.)
- Cabrera-León, Andrés, Miguel Ángel Cantero-Braojos, Llenalia Garcia-Fernandez y Juan Antonio Guerra-de-Hoyos
 2016 "Disabling chronic pain: characteristics, multimorbidity, and associated factors based on a face-to-face cross-sectional population study", *Pain Medicine (FI: 2,34; Q1; Ranking: 37/154)*, en revisión. (Citado en las pág. VIII, XII, 121.)
- Cabrera-León, Andrés, María del Mar Rueda y Miguel Ángel Cantero-Braojos
 2016 "Calibrated prevalence of disabling chronic pain", *Pain (FI=5,2; D1; Ranking: 2/30)*, en revisión. (Citado en las pág. VIII, XII, 121.)
- Cassel, CM, CE Särndal y JH Wretman
 1976 "Some results on generalized difference estimation and generalized regression estimation for finite populations", *Biometrika*, 63, pág. 615-620. (Citado en las pág. 27, 38, 55.)

Cea-D'Ancona, MA

- 2005 "La senda tortuosa de la "calidad" de la encuesta", *Revista Española de Investigaciones Sociológicas*, 111, pág. 29. (Citado en las pág. 14, 17.)

Chauvet, G, JC Deville, Tirari M y J Guennec

- 2005 *Evaluation de trois logiciels de calage: g-Calib, Calmar et Bascula*, Statistics Belgium, Working Paper, vol. 11, pág. 92. (Citado en la pág. 91.)

Chen, J y J Qin

- 1993 "Empirical likelihood estimation for finite populations and the effective usage of auxiliary information", *Biometrika*, 80, pág. 107-116. (Citado en las pág. 35-37, 40.)

Chen, J y R Sitter

- 1999 "A pseudo empirical likelihood approach to the effective use of auxiliary information in complex surveys", *Statistica Sinica*, 9, pág. 385-406. (Citado en las pág. 35, 37, 44, 119, 144.)

Chen, SX y JK Kim

- 2014 "Two-phase sampling experiment for propensity score estimation in self-selected samples", *Annals of Applied Statistics*, 8, 3, pág. 1492-1515, DOI: 10.1214/14-aos746. (Citado en la pág. 74.)

Chilet-Rosell, E, C Álvarez-Dardet y A Domingo-Salvany

- 2012 "Use of Spanish proposals for measuring social class in health sciences", *GacSanit*, 26, pág. 566-9, http://scielo.isciii.es/pdf/gs/v26n6/original_breve1.pdf. (Citado en la pág. 124.)

Costa Cabral, DM, ES Botelho Bracher, JD Prescatan Depintor y J Eluf-Neto

- 2014 "Chronic Pain Prevalence and Associated Factors in a Segment of the Population of Sao Paulo City", *J Pain*, 15, 11, pág. 1081-91. (Citado en la pág. 121.)

Dapcich, V, G Salvador, L Ribas, C Pérez, J Aranceta y Ll Serra-Majem

- 2013 *Guía de la alimentación saludable*, Sociedad Española de Nutrición Comunitaria, http://aesan.msssi.gob.es/AESAN/docs/docs/publicaciones_estudios/nutricion/guia_alimentacion.pdf. (Citado en la pág. 124.)

- De Graaff, A. A., J. A. Land, A. G. Kessels y J. L. Evers
 2011 "Demographic age shift toward later conception results in an increased age in the subfertile population and an increased demand for medical care", en *Fertil Steril*, 2011 American Society for Reproductive Medicine. Published by Elsevier Inc, United States, vol. 95, pág. 61-3, DOI: [10.1016/j.fertnstert.2010.05.013](https://doi.org/10.1016/j.fertnstert.2010.05.013). (Citado en las pág. [117](#), [119](#).)
- Deville, JC
 2005 "Calibration: past, present and future?", *Paper presented at the Workshop on "Calibration tools for survey statisticians"*, Neuchâtel, 8-9 September. (Citado en la pág. [74](#).)
- Deville, JC y CE Särndal
 1992 "Calibration estimators in survey sampling", *Journal of the American Statistical Association*, 87, pág. 376-382. (Citado en las pág. [53](#), [55](#), [56](#), [59](#), [60](#), [69](#), [86](#), [90](#).)
- Deville, JC, CE Särndal y O Sautory
 1993 "Generalized raking procedures in survey sampling", *Journal of the American Statistical Association*, 88, pág. 1013-1020. (Citado en la pág. [96](#).)
- Dobson, Annette J., Richard Hockey, Wendy J. Brown, Julie E. Byles, Deborah J. Loxton, Deidre P. McLaughlin, Leigh R. Tooth y Gita D. Mishra
 2015 "Cohort Profile Update: Australian Longitudinal Study on Women's Health", *International Journal of Epidemiology*, DOI: [10.1093/ije/dyv110](https://doi.org/10.1093/ije/dyv110), <http://ije.oxfordjournals.org/content/early/2015/06/29/ije.dyv110.abstract>. (Citado en la pág. [3](#).)
- Domar, A. D., P. C. Zuttermeister y R. Friedman
 1993 "The psychological impact of infertility: a comparison with patients with other medical conditions", *J Psychosom Obstet Gynaecol*, 14 Suppl, pág. 45-52, <http://www.ncbi.nlm.nih.gov/pubmed/8142988>. (Citado en la pág. [117](#).)
- Dominick, CH, FM Blyth y MK Nicholas
 2014 "Unpacking the burden: understanding the relationships between chronic pain and comorbidity in the general population", *Pain*, 153, 2, pág. 293-304. (Citado en las pág. [122](#), [142](#), [143](#).)

- Dueñas, M, A Salazar, B Ojeda, F Fernández-Palacin, JA Micó, LM Torres e I Failde
2014 "A Nationwide study of chronic pain prevalence in the general Spanish population: identifying clinical subgroups through cluster analysis", *Pain Med*, 16, 4, pág. 811-22. (Citado en las pág. 121, 140, 141.)
- Dupont, F
1995 "Alternative adjustments where there are several levels of auxiliary information", *Survey Methodology*, 21, pág. 125-135. (Citado en la pág. 73.)
- Dykes, L, S Singh, S Sedory y V Louis
2015 "Calibrated estimators of population mean for a mail survey design", *Communications in Statistics - Theory and Methods*, (in press), DOI: 10.1080/03610926.2013.841932. (Citado en la pág. 75.)
- Eriksen, J, MK Jensen, P Sjogren, O Ekholm y NK Rasmussen
2003 "Epidemiology of chronic non-malignant pain in Denmark", *Pain*, 106, 2, pág. 221-8. (Citado en las pág. 122, 142.)
- Estevao, V, MA Hidiroglou y CE Särndal
1995 "Methodological principles for a generalized estimation system at Statistics Canada", *Journal of Official Statistics*, 11, pág. 181-204. (Citado en la pág. 91.)
- Estevao, VM y CE Särndal
2006 "Survey Estimates by Calibration on Complex Auxiliary Information", *International Statistical Review*, 74, pág. 127-147. (Citado en las pág. 61, 144.)
- Eurostat
2011 *European Statistics Code of Practice*, Government Document, DOI: 10.2785/18474, <http://ec.europa.eu/eurostat/quality>. (Citado en la pág. 6.)
2013a *ESS Handbook for Quality Reports*, Government Document, <http://ec.europa.eu/eurostat/web/quality/quality-reporting>. (Citado en la pág. 3.)
2013b *European Health Interview Survey (EHIS wave 2) - Methodological manual*, Government Document, DOI: 10.2785/43280, <http://ec.europa.eu/eurostat/documents/3859598/5926729/KS-RA-13-018-EN.PDF/26c7ea80-01d8-420e-bdc6-e9d5f6578e7c>. (Citado en la pág. 3.)

- Evers, AWM, FW Kraaimat, R Geenen, JWG Jacobs y JWJ Bijlsma
 2003 "Pain coping and social support as predictors of long-term functional disability and pain in early rheumatoid arthritis", *Behav Res Ther*, 41, pág. 1295-1310. (Citado en la pág. 142.)
- Ewing, JA
 1984 "Detecting alcoholism. The CAGE questionnaire", *JAMA*, 252, pág. 1905-7. (Citado en la pág. 124.)
- Falorsi, PD y S Falorsi
 1997 "The Italian Generalised Package for Weighting Persons and Families: Some Experimental Results with Different Non-Response Models", *Statistics in Transition*, 3, pág. 357-381. (Citado en la pág. 95.)
- Fan, J e I Gijbels
 1996 *Local Polynomial Modelling and its Applications*, Chapman & Hall, London. (Citado en las pág. 45, 48.)
- Fernández Cordón, Juan Antonio y Constanza Tobío
 2007 *ANDALUCÍA: dependencia y solidaridad en las redes familiares*, Government Document, <http://www.juntadeandalucia.es/institutodeestadisticaycartografia./redesfamiliares/redesFamiliares.pdf>. (Citado en la pág. 3.)
- Fienberg, SE
 2011 "Bayesian Models and Methods in Public Policy and Government Settings", *Statistical Science*, 26, pág. 212-226. (Citado en la pág. 74.)
- Firth, D y KE Bennett
 1998 "Robust models in probability sampling", *J. R. Stat. Soc.*, 60, pág. 3-21. (Citado en la pág. 73.)
- Fox, K, MA Garcia, D Ardissino, P Buszman, PG Camici, F Crea, C Daly, G De Backer, P Hjerdahl, J Lopez-Sendon, J Marco, J Morais, J Pepper, U Sechtem, M Simoons, K Thygesen, SG Priori, JJ Blanc, A Budaj, J Camm, V Dean, J Deckers, K Dickstein, J Lekakis, K McGregor, M Metra, J Morais, A Osterspey, J Tamargo y JL Zamorano
 2006 "Guidelines of management of stable Angina Pectoris: the task force on the management of stable angina pectoris of the European Society of Cardiology", *Eur Heart J*, 27, pág. 1341-81. (Citado en la pág. 142.)

- Franz-Vasdeki, J, BA Pratt, M Newsome y S Germann
 2014 "Taking mHealth solutions to scale: enabling environments and successful implementation", *J Mob Technol Med*, 4(1), pág. 4, DOI: [10.7309/jmtm.4.1.8](https://doi.org/10.7309/jmtm.4.1.8). (Citado en la pág. 9.)
- Fuentes, A. y L. Devoto
 1994 "Infertility after 8 years of marriage: a pilot study", *Hum Reprod*, 9, 2, pág. 273-8, <http://www.ncbi.nlm.nih.gov/pubmed/8027283>. (Citado en las págs. 103, 104.)
- Fuller, WA y LF Burmeister
 1972 "Estimation for samples selected from two overlapping frames", *Proceedings of social science section of The American Statistical Association*, pág. 245-249. (Citado en las págs. 84, 85, 92.)
- García-Fernández, F y JA Mayor-Gallego
 1995 *Muestreo en poblaciones finitas: curso básico*, pág. 350. (Citado en la pág. 34.)
- Gelman, A, JB Carlin, HS Stern, DB Dunson, A Vehtari y DB Rubin
 2013 *Bayesian Data Analysis. 3rd edition*, London: Chapman y Hall/CRC. (Citado en la pág. 74.)
- Gerrits, MM, HW van Marwijk, P van Oppen, H van der Horst y BW Penninx
 2015 "Longitudinal association between pain, and depression and anxiety over four years", *J Psychosom Res*, 78, 1, pág. 64-70. (Citado en la pág. 143.)
- Gibson, SJ y MD Lussier
 2012 "Prevalence and relevance of pain in older persons", *Pain Med*, 13, 2, pág. 23-6. (Citado en la pág. 141.)
- Gnoth, C., E. Godehardt, P. Frank-Herrmann, K. Friol, J. Tigges y G. Freundl
 2005 "Definition and prevalence of subfertility and infertility", *Human Reproduction*, 20, 5, pág. 1144-1147, <http://humrep.oxfordjournals.org/content/20/5/1144.abstract>. (Citado en la pág. 104.)
- Godambe, VP
 1955 "A unified theory of sampling from finite populations", *Journal of the Royal Statistical Society*, 17, pág. 269-278. (Citado en las págs. 25, 38, 64.)

Godambe, VP y ME Thompson

- 1973 "Estimation in sampling theory with exchangeable prior distributions", *Annals of Statistics*, 1, 6, Godambe, vp thompson, me, pág. 1212-1221, DOI: [10.1214/aos/1176342571](https://doi.org/10.1214/aos/1176342571), %3CGo%20to%20ISI%3E://WOS:A1973R380100018. (Citado en la pág. 38.)

Goedeme, T.

- 2013 "How much Confidence can we have in EU-SILC? Complex Sample Designs and the Standard Error of the Europe 2020 Poverty Indicators", *Social Indicators Research*, 110, 1, pág. 89-110, DOI: [10.1007/s11205-011-9918-2](https://doi.org/10.1007/s11205-011-9918-2), %3CGo%20to%20ISI%3E://WOS:000314336300006. (Citado en la pág. 4.)

Gottgens, R, V Bert, O Marly y H Lon

- 1991 "Bascula, version 1.0 / A weighting package under MS-DOS. User Manual". (Citado en la pág. 89.)

Greil, A. L., J. McQuillan, K. Johnson, K. Slauson-Blevins y K. M. Shreffler

- 2010 "The hidden infertile: infertile women without pregnancy intent in the United States", *Fertil Steril*, 93, 6, pág. 2080-3, DOI: [10.1016/j.fertnstert.2009.08.024](https://doi.org/10.1016/j.fertnstert.2009.08.024), <http://www.ncbi.nlm.nih.gov/pubmed/19782354>. (Citado en la pág. 119.)

Groves, Dillman, Eltinge y Little

- 2001 *Survey Nonresponse*, New York: Wiley. (Citado en la pág. 15.)

Gutiérrez, HA

- 2015 *R package TeachingSampling: Selection of Samples and Parameter Estimation in Finite Population*, pág. 89, <https://cran.r-project.org/web/packages/TeachingSampling/index.html>. (Citado en la pág. 98.)

Guttmacher, A. F.

- 1956 "Factors affecting normal expectancy of conception", *J Am Med Assoc*, 161, 9, pág. 855-60, <http://www.ncbi.nlm.nih.gov/pubmed/13319020>. (Citado en la pág. 103.)

Guzi, M y PD Garcia

- 2015 "A web survey analysis of subjective well-being", *International Journal of Manpower*, 36, 1, pág. 48-67, DOI: [10.1108/ijm-12-2014-0237](https://doi.org/10.1108/ijm-12-2014-0237). (Citado en la pág. 74.)

Härdle, W y JS Marron

- 1991 "Bootstrap Simultaneous Error Bars for Nonparametric Regression", 19, 2, pág. 778-796. (Citado en la pág. 45.)

Harms, T

- 2003 *Calibration estimators for prediction of dynamics in panels. Using longitudinal patterns to improve calibration estimates about developments in panels*, Wiesbaden (Germany): Statistisches Bundesamt, pág. 22, <http://hdl.handle.net/10068/91524>. (Citado en la pág. 74.)

Harms, T y P Duchesne

- 2006 "On calibration estimation for quantiles", *Survey Methodology*, 32, pág. 37-52. (Citado en la pág. 70.)

Hartley, HO

- 1962 "Multiple frame surveys", en *Proceedings of the Social Statistics Section, American Statistical Association*, pág. 203-206. (Citado en las págs. 82, 84, 92.)

Hartley, HO y JNK Rao

- 1968 "A NEW ESTIMATION THEORY FOR SAMPLE SURVEYS", *Biometrika*, 55, 3, Hartley, ho rao, jnk, pág. 547-599, DOI: [10.2307/2334260](https://doi.org/10.2307/2334260), [%3CGo%20to%20ISI%3E://WOS:A1968C214100011](https://www.crossref.org/wos/A1968C214100011). (Citado en la pág. 35.)

Hartley, HO y A Ross

- 1954 "UNBIASED RATIO ESTIMATORS", *Nature*, 174, 4423, Hartley, ho ross, a, pág. 270-271, DOI: [10.1038/174270a0](https://doi.org/10.1038/174270a0), [%3CGo%20to%20ISI%3E://WOS:A1954UA39700025](https://www.crossref.org/wos/A1954UA39700025). (Citado en la pág. 32.)

Hedayat, AS y BK Sinha

- 1991 *Design and inference infinite population sampling*. (Citado en la pág. 27.)

Herrera-Silva, J, J Rodríguez-Matallana, D Contreras de la Fuente, R de la Torre-Liébanas, F Gómez-Armenta, F Linares del Río, I López-Moraleda, F Neira Reina, M Rodríguez López, R Rodríguez de la Torre, A Rubio Andrada y S Ruiz Ortiz

- 2012 "Epidemiological study of pain in Andalusia (Spain)", *Rev SocEsp Dolor*, 19, 5, pág. 239-51. (Citado en las págs. 121, 141.)

Hill, BM y A Shaw

- 2013 "The Wikipedia Gender Gap Revisited: Characterizing Survey Response Bias with Propensity Score Estimation", *Plos One*, 8, 6, DOI: [10.1371/journal.pone.0065782](https://doi.org/10.1371/journal.pone.0065782). (Citado en la pág. 74.)

Horvitz, DG y DJ Thompson

- 1953 "A generalization of sampling without replacement from a finite universe", *Journal of the American Statistical Association*, 47, pág. 663-685. (Citado en las págs. 29, 53, 54, 80.)

INE

- 1999 *The national survey of fecundity*, Government Document, <http://www.ine.es/jaxi/menu.do?type=pcaxis&path=/t20/p317&file=inebase&L=0>. (Citado en la pág. 104.)
- 2009 *Panorámica de la discapacidad en España*, Government Document, <http://www.ine.es/jaxi/menu.do?type=pcaxis&path=/t15/p418&file=inebase&L=0>. (Citado en la pág. 3.)
- 2010 *Cifras Oficiales de Población de los Municipios Españoles: Revisión del Padrón municipal a 1 de enero de 2010*, National Statistics Institute, <http://www.ine.es/jaxi/menu.do?type=pcaxis&path=/t20/e260&file=inebase&L=0>. (Citado en las pág. 105, 106.)
- 2011 *Metodología de los Censos de Población y Viviendas 2011*, http://www.ine.es/censos2011_datos/cen11_datos_metodologia.htm. (Citado en las pág. 14, 126, 127.)
- 2015a *Encuesta sobre equipamiento y uso de tecnologías de información y comunicación en los hogares*, Government Document, http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176741&menu=ultiDatos&idp=1254735976608. (Citado en la pág. 77.)
- 2015b *Population and Housing Census 2011*, National Statistics Institute, http://www.ine.es/en/censos2011_datos/cen11_datos_inicio_en.htm. (Citado en las pág. 107, 113, 124, 138, 139.)
- Jausoro, A
- 2000 *Reproducción Humana Asistida: Descripción de las opciones terapéuticas disponibles*, Generic. (Citado en la pág. 104.)
- Joffe, M., J. Key, N. Best, T. K. Jensen y N. Keiding
- 2009 "The role of biological fertility in predicting family size", *Hum Reprod*, 24, 8, pág. 1999-2006, DOI: 10.1093/humrep/dep087, <http://www.ncbi.nlm.nih.gov/pubmed/19429909>. (Citado en la pág. 119.)
- Johannes, CB, TK Le, X Zhou, JA Johnston y RH Dworkin
- 2010 "The prevalence of chronic pain in United States adults: results of an Internet-based survey", *J Pain*, 11, 11, pág. 1230-39. (Citado en las pág. 121, 140, 141.)
- Kalton, G y DW Anderson
- 1986 "Sampling rare populations", *Journal of the Royal Statistical Society. Series A (General)*, 149, pág. 65-82. (Citado en la pág. 86.)

Karmaus, W y S Juul

- 1999 "Infertility and subfecundity in population-based samples from Denmark, Germany, Italy, Poland and Spain", *Eur J Public Health*, 9, pág. 229-235. (Citado en las pág. [105](#), [118](#).)

Kennedy, J, JM Roll, T Schraudner, S Murphy y S McPherson

- 2014 "Prevalence of persistent pain in the US adult population: new data from the National Health Interview Survey", *J Pain*, 15, 10, pág. 979-84. (Citado en las pág. [121](#), [141](#).)

Kerns, RD, R Rosenberg y JD Otis

- 2002 "Self-appraised problem solving and pain-relevant social support as predictors of the experience of chronic pain", *Ann Behav Med*, 24, pág. 100-105. (Citado en la pág. [141](#).)

Kott, Phillip S

- 2009 "Calibration Weighting: Combining Probability Samples and Linear Prediction Models", en *SAMPLE SURVEYS: INFERENCE AND ANALYSIS, VOL 29B*, ed. por Pfeffermann, D and Rao, CR, Handbook of Statistics, ELSEVIER SCIENCE BV, vol. 29B, 55-82, DOI: [{10.1016/S0169-7161\(09\)00225-9}](#). (Citado en la pág. [81](#).)
- 2008 *Building a Better Delete-a-Group Jackknife for a Calibration Estimator*, NASS Research Report, Washington. (Citado en la pág. [92](#).)

Kott, PS y CD Day

- 2014 "Developing Calibration Weights and Standard-Error Estimates for a Survey of Drug-Related Emergency-Department Visits", *Journal of Official Statistics*, 30, 3, pág. 521-532, DOI: [dx.doi.org/10.2478/JOS-2014-0032](#). (Citado en la pág. [144](#).)

Kovacevic, M

- 1997 "Calibration estimation of cumulative distribution and quantile functions from survey data", *Proceedings of the Survey Methods Section, Statistical Society of Canada*, 47, pág. 139-144. (Citado en la pág. [71](#).)

Kuo, L

- 1988 "Classical and Prediction Approaches to Estimating Distribution Functions From Survey Data", en *ASA Proceedings of the Section on Survey Research Methods*, American Statistical Association (Alexandria, VA), pág. 280-285. (Citado en las pág. [45](#), [46](#).)

Langley, PC, MA Ruiz-Iban, JT Tornero-Molina, J de-Andres y JR González-Escalada-Castellón

- 2011 "The prevalence, correlates and treatment of pain in Spain", *J. Media Econ*, 14, 3, pág. 367-80. (Citado en las pág. [121](#), [140](#), [141](#).)

Larsen, U.

- 2005 "Research on infertility: which definition should we use?", en *Fertil Steril*, United States, vol. 83, pág. 846-52, DOI: [10.1016/j.fertnstert.2004.11.033](#). (Citado en la pág. [103](#).)

Le Guennec, J y O Sautory

- 2004 "CALMAR 2: Une nouvelle version de la macro Calmar de redressement d échantillon par calage", *Echantillonnage et méthodes d enquêtes*, pág. 375. (Citado en la pág. [90](#).)

Lee, S y R Valliant

- 2009 "Estimation for Volunteer Panel Web Surveys Using Propensity Score Adjustment and Calibration Adjustment", *Sociological Methods Research*, 37, 3, pág. 319-343, DOI: [10.1177/0049124108329643](#). (Citado en la pág. [74](#).)

Legleye, S, G Charrance, N Razafindratsima, N Bajos, A Bohet y C Moreau

- 2015 "Estimation for Volunteer Panel Web Surveys Using Propensity Score Adjustment and Calibration Adjustment", *Sociological Methods Research*, DOI: [10.1177/0049124115621333](#). (Citado en las pág. [74](#), [145](#).)

Leonardi, M, J Bickenbach, TB Ustun, N Kostanjsek y S Chatterji

- 2006 "The definition of disability: what is in a name?", *Lancet*, 368, pág. 1219-21. (Citado en la pág. [141](#).)

Little, RJ

- 2003 "The Bayesian Approach to Sample Survey Inference", en *Analysis of Survey Data*, ed. por Chambers y CJ Skinner, New York: Wiley, pág. 49-57. (Citado en la pág. [74](#).)
- 2012 "Calibrated Bayes, an Alternative Inferential Paradigm for Official Statistics", *Journal of Official Statistics*, 28, pág. 309-372. (Citado en la pág. [74](#).)
- 2015 "Calibrated Bayes, an inferential paradigm for official statistics in the era of big data", *Statistical Journal of the IAOS*, 31, pág. 555-563, DOI: [10.3233/SJI-150944](#). (Citado en la pág. [75](#).)

Little, RJ y DB Rubin

- 2016 *Missing Data*, New York: Wiley. (Citado en las pág. [15](#), [74](#).)

Little, RJA

- 2004 "To Model or Not to Model? Competing Modes of Inference for Finite Population Sampling", *Journal of the American Statistical Association*, 99, 466, pág. 546-556. (Citado en la pág. 46.)

Lohr, SL

- 2009 "Multiple-frame surveys", *Handbook of Statistics*, 29, pág. 71-88. (Citado en las pág. 83, 84.)

Lopez Escobar, Emilio y Ernesto Barrios Zamudio

- 2015 *samplingVarEst: Sampling Variance Estimation*, <https://cran.r-project.org/web/packages/samplingVarEst/samplingVarEst.pdf>. (Citado en las pág. 111, 127.)

López, V

- 2014 "Epidemiology of infertility. The impact of environmental factors", *Rev Iberoam Fert Rep Hum*, 31, pág. 31-37. (Citado en la pág. 119.)

Lumley, T

- 2015 *R Package survey: analysis of complex survey samples*, pág. 129, <https://cran.r-project.org/web/packages/survey/survey.pdf>. (Citado en las pág. 96, 98.)

Lumsdaine, Robin L. y Anneke Exterkate

- 2013 "How survey design affects self-assessed health responses in the Survey of Health, Ageing, and Retirement in Europe (SHARE)", *European Economic Review*, 63, pág. 299-307, DOI: [10.1016/j.euroecorev.2013.06.002](https://doi.org/10.1016/j.euroecorev.2013.06.002), %3CGo%20to%20ISI%3E://WOS:000324974500019. (Citado en la pág. 4.)

Lundström, S y CE Särndal

- 1999 "Calibration as a standard method for treatment of nonresponse", *Journal of Official Statistics*, 15, pág. 305-327. (Citado en la pág. 74.)
- 2001 *Estimation in the Presence of Nonresponse and Frame Imperfections*, Statistics Sweden, pág. 172. (Citado en las pág. 71, 122.)

Lynn, P, N Buck, J Burton, A Jäckle y H Laurie

- 2005 *A review of methodological research pertinent to longitudinal survey design and data collection*, Working Paper. (Citado en la pág. 6.)

Martínez, Sergio, Antonio Arcos, Helena Martínez y Sarjinder Singh

- 2015 "Estimating Population Proportions by Means of Calibration Estimators", *Revista Colombiana de Estadística*, 38, 1, pág. 267-293, <http://www.revistas.unal.edu.co/index.php/estad/article/view/48814>. (Citado en la pág. 75.)

- Martínez, Sergio, María del Mar Rueda, Antonio Arcos y H Martínez
 2010 "Optimum calibration points estimating distribution functions", *Journal of Computational and Applied Mathematics*, 233, 9, pág. 2265-2277. (Citado en la pág. 71.)
- Martínez, Sergio, María del Mar Rueda, Antonio Arcos, H Martínez y JF Muñoz
 2012 "On determining the calibration equations to construct model-calibration estimators of the distribution function", *Revista Matemática Complutense*, 25, 1, pág. 87-95, DOI: [10.1007/s13163-010-0058-z](https://doi.org/10.1007/s13163-010-0058-z). (Citado en la pág. 71.)
- Martínez, Sergio, María del Mar Rueda, H Martínez y Antonio Arcos
 2015 "Determining P optimum calibration points to construct calibration estimators of the distribution function", *Journal of Computational and Applied Mathematics*, 275, pág. 281-293. (Citado en la pág. 71.)
- Maruani, P. y D. Schwartz
 1983 "Sterility and fecundability estimation", *J Theor Biol*, 105, 2, pág. 211-9, <http://www.ncbi.nlm.nih.gov/pubmed/6656280>. (Citado en la pág. 103.)
- Matorras Weinig, RJ
 2011 *Libro blanco sociosanitario: la infertilidad en España, situación actual y perspectivas*, Imago Concept Image DEV, Madrid. (Citado en la pág. 118.)
- McEwen, BS
 1998 "Stress, adaptation, and disease: Allostasis and allostatic load", *Ann New York Acad Sci*, 840, 1, pág. 33-44. (Citado en las pág. 122, 142.)
- Menard, S
 2000 "Coefficients of determination for multiple logistic regression analysis", *The American Statistician*, 54, pág. 17-24. (Citado en la pág. 125.)
- Menken, J. y U. Larsen
 1994 "Estimating the incidence and prevalence and analyzing the correlates of infertility and sterility", *Ann N Y Acad Sci*, 709, pág. 249-65, <http://www.ncbi.nlm.nih.gov/pubmed/8154718>. (Citado en la pág. 117.)

El-Messidi, A., H. Al-Fozan, S. Lin Tan, R. Farag y T. Tulandi

- 2004 "Effects of repeated treatment failure on the quality of life of couples with infertility", *J Obstet Gynaecol Can*, 26, 4, pág. 333-6, <http://www.ncbi.nlm.nih.gov/pubmed/15115622>. (Citado en la pág. 117.)

Miettinen, OS y EF Cook

- 1981 "Confounding: Essence and detection", *American Journal of Epidemiology*, 144, 4, pág. 593-603. (Citado en la pág. 125.)

Ministerio de Sanidad

- 2013 *Presentación de la Encuesta Nacional de Salud*, Government Document, <http://www.msssi.gob.es/estadEstudios/estadisticas/encuestaNacional/encuestaNac2011/PresentacionENSE2012.pdf>. (Citado en la pág. 3.)
- 2015 *Encuestas de Salud de las comunidades autónomas en la red*, Web Page, <http://www.msssi.gob.es/estadEstudios/estadisticas/encuestaNacional/encuestasCCAA.htm>. (Citado en la pág. 3.)

Montanari, Giorgio E y M Giovanna Ranalli

- 2005 "Nonparametric model calibration estimation in survey sampling", *Journal of the American Statistical Association*, 100, 472, pág. 1429-1442. (Citado en la pág. 48.)
- 2009 "Multiple and ridge model calibration for sample surveys", en *Proceedings of the Workshop in Calibration and estimation in surveys*, Statistics Canada Ed. (Citado en la pág. 74.)

Muñoz, FJ

- 2006 *Aportaciones a los métodos de estimación de parámetros lineales y no lineales con información auxiliar*, Tesis doctoral. Universidad de Granada. (Citado en la pág. 44.)

Navarro, JL, L Martínez, JA Castilla y Hernández E

- 2006 "Cost of Assisted Reproduction Technology in a Public Hospital", *Gac Sanit*, 20, pág. 382-391. (Citado en la pág. 103.)

Nübling, M, H Burr, S Moncada y T.S Kristensen

- 2014 "COPSOQ International Network: Co-operation for research and assessment of psychosocial factors at work", *Public Health Forum*, 22, 1, 18.e1-18.e3. (Citado en la pág. 124.)

Oakley, L., P. Doyle y N. Maconochie

- 2008 "Lifetime prevalence of infertility and infertility treatment in the UK: results from a population-based survey of reproduction", *Hum Reprod*, 23, 2, pág. 447-50, DOI: [10.1093/humrep/dem369](https://doi.org/10.1093/humrep/dem369). (Citado en la pág. 113.)

Observatorio Español de Drogas

- 2013 *Informe nacional Reitox: España 2012/Resultados EDADES 2012*, Government Document, http://www.pnsd.msssi.gob.es/cendocubd/jsp/file/file_explore.jsp?recordId=default:ELEUSIS:30573. (Citado en la pág. 3.)

Ohayon, MM y AF Schatzberg

- 2010 "Chronic pain and major depressive disorder in the general population", *J Psychiat Res*, 44, 7, pág. 454-61. (Citado en la pág. 142.)

Ohayon, MM y JC Stingl

- 2012 "Prevalence and comorbidity of chronic pain in the German general population", *J Psychiat Res*, 46, 4, pág. 444-50. (Citado en la pág. 142.)

Olsen, RB, S Bruehl, ChS Nielsen, LA Rosselandb, AE Eggen y A Stubhaug

- 2013 "Hypertension prevalence and diminished blood pressure-related hypoalgesia in individuals reporting chronic pain in a general population: The Tromso Study", *Pain*, 154, 2, pág. 257-62. (Citado en la pág. 142.)

Ombelet, W., I. Cooke, S. Dyer, G. Serour y P. Devroey

- 2008 "Infertility and the provision of infertility medical services in developing countries", *Hum Reprod Update*, 14, 6, pág. 605-21, <http://www.ncbi.nlm.nih.gov/pubmed/18820005>. (Citado en la pág. 117.)

Opsomer, JD y CP Miller

- 2005 "Selecting the Amount of Smoothing in Nonparametric Regression Estimation for Complex Surveys", *Journal of Nonparametric Statistics*, 17, pág. 593-611. (Citado en la pág. 45.)

Owen, AB

- 1988 "EMPIRICAL LIKELIHOOD RATIO CONFIDENCE-INTERVALS FOR A SINGLE FUNCTIONAL", *Biometrika*, 75, 2, Owen, ab, pág. 237-249, DOI: [10.1093/biomet/75.2.237](https://doi.org/10.1093/biomet/75.2.237), %3CGo%20to%20ISI%3E://WOS:A1988N941300007. (Citado en la pág. 35.)

Owen, AB

- 1990 "EMPIRICAL LIKELIHOOD RATIO CONFIDENCE-REGIONS", *Annals of Statistics*, 18, 1, Owen, a, pág. 90-120, DOI: [10.1214/aos/1176347494](https://doi.org/10.1214/aos/1176347494), %3CGo%20to%20ISI%3E://WOS:A1990DA37500004. (Citado en la pág. 35.)
- 2001 *Empirical Likelihood*, Chapman y Hall-CRC Press. (Citado en la pág. 35.)

Park, M y WA Fuller

- 2009 "The mixed model for survey regression estimation", *J. Stat. Plann. Infer.*, 139, pág. 1320-1331. (Citado en la pág. 74.)

Peinado, J. A. y F. Bolúmar

- 1997 "[Thoughts about sterility in Spain]", *Med Clin (Barc)*, 109, 15, pág. 585-7, <http://www.ncbi.nlm.nih.gov/pubmed/9441194>. (Citado en las pág. 103, 104.)

Pentland, A, TG Reid y T Heibeck

- 2013 *Big Data and Health: Revolutionizing medicine and Public Health*, World Innovation Summit for Health, pág. 40, <http://wish-qatar.org/big-data/big-data>. (Citado en la pág. 10.)

Pérez, C y MB Portero

- 2015 "Técnicas de Calibrado dirigidas a encuestas de hogares", en *Escuela de Estadística de las Administraciones Públicas*, Instituto Nacional de Estadística, pág. 1-52. (Citado en la pág. 90.)

Pérez, RA

- 2002 "¿Qué es un modelo de superpoblación?", *Metodología de Encuestas*, 4, pág. 79-86. (Citado en la pág. 25.)

Peterson, BD

- 2006 *Examining the individual and dyadic coping processes of men and women in infertile couples and their relationship to infertility stress, marital adjustment, and depression*. Generic. (Citado en la pág. 117.)

Pfeffermann y Rao

- 2009 *Handbook of Statistics Sample Surveys: Design, Methods and Applications*, Handbook of Statistics, Pfeffermann & Rao, vol. 29A, pág. 722. (Citado en la pág. 9.)

Portenoy, RK, C Ugarte, I Fuller y G Haas

- 2004 "Population-Based Survey of pain in the United States: differences among White, African American, and Hispanic subjects", *J Pain*, 5, 6, pág. 317-28. (Citado en la pág. 143.)

Prieto Lara, E y Ocaña Riola R

- 2010 "Updating Rurality Index for small areas in Spain", *Soc Indic Res*, 89, pág. 267-280. (Citado en la pág. 124.)

Qin, J y J Lawless

- 1994 "EMPIRICAL LIKELIHOOD AND GENERAL ESTIMATING EQUATIONS", *Annals of Statistics*, 22, 1, pág. 300-325, DOI: 10.1214/aos/1176325370, %3CGo%20to%20ISI%3E://WOS:A1994NH41200014. (Citado en la pág. 35.)

- 1995 "ESTIMATING EQUATIONS, EMPIRICAL LIKELIHOOD AND CONSTRAINTS ON PARAMETERS", *Canadian Journal of Statistics- Revue Canadienne De Statistique*, 23, 2, Qin, j lawless, j, pág. 145-159, DOI: 10.2307/3315441, %3CGo%20to%20ISI%3E://WOS:A1995RN85400004. (Citado en la pág. 35.)

R Core Team

- 2015 *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org/>. (Citado en las pág. 12, 89, 92, 111, 127.)

Raftery, MN, K Sarma, AW Murphy, D De la Harpe, C Normand y BE McGuire

- 2011 "Chronic pain in the Republic of Ireland-community prevalence, psychosocial profile and predictors of pain-related disability: results from the Prevalence, Impact and Cost of Chronic Pain (PRIME) study, part 1", *Pain*, 152, 5, pág. 1096-103. (Citado en las pág. 121, 122, 142.)

Ranalli, M Giovanna, Antonio Arcos, María del Mar Rueda y Annalisa Teodoro

- 2015 "Calibration estimation in dual-frame surveys", *Statistical Methods and Applications*, First online: 01 September 2015, pág. 1-29. (Citado en las pág. 86, 87.)

Rao, JNK

- 1994 "Estimating totals and distribution functions using auxiliary information at the estimation stage", *Journal of Official Statistics*, 10, 2, pág. 153. (Citado en las pág. 33, 34.)

Rao, JNK y C Wu

- 2010 "Pseudo-Empirical Likelihood Inference for Multiple Frame Surveys", *Journal of the American Statistical Association*, 105, 492, pág. 1494-1503. (Citado en las pág. 85, 92.)

Reid, MC, C Eccleston y K Pillemer

- 2015 "Management of chronic pain in older adults", *BMJ*, 350, pág. 532. (Citado en las pág. 121, 122, 142.)

Rueda, María del Mar, Sergio Martínez, H Martínez y Antonio Arcos

- 2007 "Estimation of the distribution function with calibration methods", *Journal of Statistical Planning and Inference*, 137, 2, pág. 435-448. (Citado en la pág. 70.)

Rueda, María del Mar e I Sánchez-Borrego

- 2009 "A predictive estimator of finite population mean using non-parametric regression", *Computational Statistics*, 24, pág. 1-14. (Citado en la pág. 47.)

Rueda, María del Mar, I Sánchez-Borrego y Gonzalez A

- 2008 "A new method for solving the problem of the mean estimation when the underlying regression function is discontinuous", *International Journal of Computer Mathematics*, 857, pág. 1073-1082. (Citado en la pág. 48.)

Sakakibara, Toshihiko, Zhuo Wang, Permsak Paholpak, Weerachai Kosuwon, T Myint y Yuichi Kasai

- 2013 "A Comparison of Chronic Pain Prevalence in Japan, Thailand, and Myanmar", *Pain Phys*, 16, pág. 603-08. (Citado en la pág. 121.)

Sánchez Cruz, José Juan, Llenalia García Fernández y José M^a Mayoral Cortés

- 2013 *La Salud en Andalucía Según las Encuestas Andaluzas de Salud (EAS): Encuesta Andaluza de Salud 2011-2012*, Escuela Andaluza de Salud Pública, pág. 347, http://www.juntadeandalucia.es/salud/sites/csalud/contenidos/Informacion_General/c_3_c_1_vida_sana/vigilancia_de_la_salud/encuesta?perfil=org. (Citado en las pág. 3, 123.)

Sánchez-Borrego, I, JD Opsomer, M Rueda y A Arcos

- 2014 "Nonparametric estimation with mixed data types in survey sampling", *Revista Matemática Complutense*, 27, 2, pág. 685-700, DOI: 10.1007/s13163-013-0142-2, <http://www.scopus.com/inward/record.url?eid=2-s2.0-84904387628&partnerID=40&md5=69d60658eaec438aa5adb37a1deffb37>. (Citado en la pág. 49.)

- Sánchez-Borrego, IR, MD Martínez-Miranda y A González-Carmona
 2006 "Local linear kernel estimation of the discontinuous regression function", *Computational Statistics*, 21, pág. 557-569. (Citado en la pág. 48.)
- Sánchez-Cantalejo, C, R OcañaRiola y FernándezAjuria A
 2008 "Deprivation index for small areas in Spain", *Soc Indic Res*, 89, pág. 259-273. (Citado en la pág. 124.)
- Sánchez-Crespo, Gonzalo
 2002 "Introducción a los modelos de superpoblación en las técnicas de muestreo con probabilidad desiguales", *Metodología de Encuestas*, 4, pág. 87-104. (Citado en la pág. 25.)
- Särndal, CE
 2007 "The calibration approach in survey: theory and practice", *Survey Methodology*, 33, 2, pág. 99-119. (Citado en las pág. 107, 119, 122.)
- Särndal, CE y VM Estevao
 2000 "A Functional Form Approach to Calibration", *Journal of Official Statistics*, 16, pág. 379-399. (Citado en la pág. 61.)
- Särndal, CE y S Lundström
 2005 *Estimation in Surveys with Nonresponse*, New York: Wiley, pág. 212. (Citado en las pág. 15, 74, 107, 119, 144.)
 2008 "Assessing auxiliary vectors for control of nonresponse bias in the calibration estimator", *Journal of Official Statistics*, 24, pág. 251-260. (Citado en la pág. 74.)
 2009 "Design for estimation: Identifying auxiliary vectors to reduce nonresponse bias", *Statistics Sweden*, pág. 60. (Citado en la pág. 74.)
- Särndal, CE, B Swensson y J Wretman
 1992 "Model Assisted Survey Sampling". (Citado en la pág. 34.)
- Särndal, CE, B Swensson y JH Wretman
 1989 "The Weighted Technique for Estimating the Variance of the General Regression Estimator of the finite population total", *Biometrika*, 76, pág. 527-537. (Citado en las pág. 46, 56.)
- Sautory, O
 1991 "La macro SAS: CALMAR (redressement d un echantillon par calage sur marges)". (Citado en la pág. 90.)

Schmidt, L., U. Christensen y B. E. Holstein

2005 "The social epidemiology of coping with infertility", *Human Reproduction*, 20, 4, pág. 1044-1052, <http://humrep.oxfordjournals.org/content/20/4/1044.abstract>. (Citado en la pág. 103.)

Schmidt, L., K. Münster y P. Helm

1995 "Infertility and the seeking of infertility treatment in a representative population", *Br J Obstet Gynaecol*, 102, 12, pág. 978-84, <http://www.ncbi.nlm.nih.gov/pubmed/8652489>. (Citado en la pág. 118.)

Schober, MF, J Pasek, L Guggenheim, C Lampe y FG Conrad

2016 "Social Media Analyses for Social Measurement", *Public Opinion Quarterly*, 80, 1, pág. 180-211, DOI: 10.1093/poq/nfv048, <http://poq.oxfordjournals.org/content/80/1/180.abstract>. (Citado en la pág. 74.)

Sexton, M. B., M. R. Byrd y S. von Kluge

2010 "Measuring resilience in women experiencing infertility using the CD-RISC: examining infertility-related stress, general distress, and coping styles", *J Psychiatr Res*, 44, 4, pág. 236-41, DOI: 10.1016/j.jpsychires.2009.06.007, <http://www.ncbi.nlm.nih.gov/pubmed/19665730>. (Citado en la pág. 103.)

Singh, HP y MR Espejo

2003 "On linear regression and ratio-product estimation of a finite population mean", *Journal of the Royal Statistical Society Series D-the Statistician*, 52, pág. 59-67, DOI: 10.1111/1467-9884.00341, %3CGo%20to%20ISI%3E://WOS:000181800800005. (Citado en las pág. 33, 34.)

Singh, S

2001 "Generalized calibration approach for estimating variance in survey sampling", *Annals of the Institute of Statistical Mathematics*, 53, pág. 404-417. (Citado en la pág. 74.)

2013 "A dual problem of calibration of design weights", *Statistics*, 47, 3, pág. 566-574. (Citado en la pág. 75.)

Singh, S, S Horn, S Chowdhury y F Yu

1999 *Calibration of the Estimators of Variance*, Australian Statistical Publishing Association, vol. 60. (Citado en la pág. 74.)

Singh, Sarjinder

2003 *Advanced Sampling Theory With Applications*, Springer Science & Business Media, vol. 2. (Citado en la pág. 125.)

- Skinner, CJ
 1991 "On the efficiency of raking ratio estimation for multiple frame surveys", *Journal of the American Statistical Association*, 86, pág. 779-784. (Citado en las pág. 86, 92.)
- Skinner, CJ y JNK Rao
 1996 "Estimation in dual frame surveys with complex designs", *Journal of the American Statistical Association*, 91, pág. 349-356. (Citado en las pág. 84, 92.)
- Slama, R., O. K. Hansen, B. Ducot, A. Bohet, D. Sorensen, L. Giorgis Allemand, M. J. Eijkemans, L. Rosetta, J. C. Thalabard, N. Keiding y J. Bouyer
 2012 "Estimation of the frequency of involuntary infertility on a nation-wide basis", *Hum Reprod*, 27, 5, pág. 1489-98, DOI: [10.1093/humrep/des070](https://doi.org/10.1093/humrep/des070), <http://www.ncbi.nlm.nih.gov/pubmed/22416008>. (Citado en la pág. 118.)
- Soares-Weiser, K e I Sola
 2014 *Epidemiology of chronic nonmalignant pain in Spain*, Tornero, J et al. (Citado en la pág. 141.)
- Sørensen, K., J. M. Pelikan, F. Röthlin, K. Ganahl, Z. Slonska, G. Doyle, J. Fullam, B. Kondilis, D. Agrafiotis, E. Uiters, M. Falcon, M. Mensing, K. Tchamov, S. V. Broucke, H. Brand y HLS-EU Consortium
 2015 "Health literacy in Europe: comparative results of the European health literacy survey (HLS-EU)", *Eur J Public Health*, pág. 1-6, DOI: [10.1093/eurpub/ckv043](https://doi.org/10.1093/eurpub/ckv043), <http://www.ncbi.nlm.nih.gov/pubmed/25843827>. (Citado en la pág. 3.)
- Sørensen, K., S. Van den Broucke, J. M. Pelikan, J. Fullam, G. Doyle, Z. Slonska, B. Kondilis, V. Stoffels, R. H. Osborne, H. Brand y HLS-EU Consortium
 2013 "Measuring health literacy in populations: illuminating the design and development process of the European Health Literacy Survey Questionnaire (HLS-EU-Q)", *BMC Public Health*, 13, pág. 948, DOI: [10.1186/1471-2458-13-948](https://doi.org/10.1186/1471-2458-13-948), <http://www.ncbi.nlm.nih.gov/pubmed/24112855>. (Citado en la pág. 3.)
- Sukhatme, PV, BV Sukhatme, S Sukhatme y C Asok
 1984 *Sampling theory with applications*, Indian Society of Agricultural Statistics, New Delhi & IOWA State University Press, Ames, USA. (Citado en las pág. 32, 34.)

Terava, A. N., M. Gissler, E. Hemminki y R. Luoto

- 2008 "Infertility and the use of infertility treatments in Finland: prevalence and socio-demographic determinants 1992-2004", en *Eur J Obstet Gynecol Reprod Biol*, Ireland, vol. 136, pág. 61-6, DOI: [10.1016/j.ejogrb.2007.05.009](https://doi.org/10.1016/j.ejogrb.2007.05.009). (Citado en la pág. 118.)

The ESHRE Capri Workshop Group

- 1994 "Male sterility and subfertility: guidelines for management." *Hum Reprod*, 9, 7, pág. 1260-4, <http://www.ncbi.nlm.nih.gov/pubmed/7962430>. (Citado en la pág. 103.)

Théberge, A

- 1999 "Extensions of calibration estimators in survey sampling", *Journal of the American Statistical Association*, 94, pág. 635-644. (Citado en la pág. 74.)

Thoma, M. E., A. C. McLain, J. F. Louis, R. B. King, A. C. Trumble, R. Sundaram y G. M. Buck Louis

- 2013 "Prevalence of infertility in the United States as estimated by the current duration approach and a traditional constructed approach", *Fertil Steril*, 99, 5, 1324-1331.e1, DOI: [10.1016/j.fertnstert.2012.11.037](https://doi.org/10.1016/j.fertnstert.2012.11.037), <http://www.ncbi.nlm.nih.gov/pubmed/23290741>. (Citado en las pág. 116, 118, 119.)

Thompson, ME

- 1988 *Superpopulation Models*, Encyclopedia of Statistical Sciences, pág. 93-99. (Citado en la pág. 25.)

Tillé, Y y A Matei

- 2015 *R Package sampling: survey sampling*, pág. 76, <https://cran.r-project.org/web/packages/sampling/sampling.pdf>. (Citado en las pág. 95, 96.)

Toth, C, J Lander y S Wiebe

- 2009 "The prevalence and impact of chronic pain with neuropathic pain symptoms in the general population", *Pain Med*, 10, 5, pág. 918-29. (Citado en la pág. 121.)

Tsang, A, M Von Korff, S Lee, J Alonso, E Karam, MC Angermeyer, GL Borges, EJ Bromet, GR di Girolamo de Graaf, O Gureje, JP Lepine, JM Haro, D Levinson, MA Oakley Browne, J Posada-Villa, S Seedat y M Watanabe

- 2008 "Common chronic pain conditions in developed and developing countries: gender and age differences and comorbidity with depression-anxiety disorders", *J Pain*, 10, 9, pág. 883-91. (Citado en la pág. 142.)

- Valliant, R, AH Dorfman y RM Royall
 2000 *Finite Population Sampling and Inference: A Prediction Approach*, Wiley, New York. (Citado en las pág. 26, 46.)
- Vanderhoeft, C
 2001 *Generalised Calibration at Statistics Belgium. SPSS Module g-CALIB-S and Current Practices*, 3, Statistics Belgium Working Paper, pág. 192. (Citado en la pág. 91.)
- Vayena, E, PJ Rowe y PD Griffin
 s.f. *Medical, ethical & social aspects of assisted reproduction. Current practices & controversies in assisted reproduction: Report of a WHO meeting, 2001*. Government Document. (Citado en la pág. 117.)
- Verbrugge, LM
 1997 "A global disability indicator", *J Aging Stud*, 11, pág. 337-62. (Citado en la pág. 124.)
- Vilagut, G, JM Valderas, M Ferrer, O Garin, E López-García y J Alonso
 2008 "Interpretation of SF-36 and SF-12 questionnaires in Spain: physical and mental components", *Medicina Clínica*, 130, 19, pág. 726-35. (Citado en la pág. 124.)
- Von Korff, M
 1992 "Grading the severity of chronic pain", *Pain*, 50, 2, pág. 133-149. (Citado en la pág. 143.)
- Wand, MP y MC Jones
 1995 *Kernel Smoothing*, Chapman y Hall, London. (Citado en la pág. 45.)
- Wang, X., C. Chen, L. Wang, D. Chen, W. Guang y J. French
 2003 "Conception, early pregnancy loss, and time to clinical pregnancy: a population-based prospective study", *Fertil Steril*, 79, 3, pág. 577-84, <http://www.ncbi.nlm.nih.gov/pubmed/12620443>. (Citado en la pág. 118.)
- Welsh, AH y E Ronchetti
 1998 "Bias-calibrated estimation from sample surveys containing outliers", *Journal of the Royal Statistical Society*, 60, pág. 413-428. (Citado en la pág. 73.)
- WHO
 2006 "International classification of functioning, disability, and health". (Citado en la pág. 121.)

WHO

- 2015a *A practical guide for engaging with mobile operators in mHealth for reproductive, maternal, newborn and child health*. Government Document, <http://www.who.int/reproductivehealth/publications/mhealth/mobile-operators-mhealth/en/>. (Citado en la pág. 9.)
- 2015b *Body Mass Classification*, World Health Organization, http://apps.who.int/bmi/index.jsp?introPage=intro_3.html. (Citado en la pág. 124.)

Wilkes, S., D. J. Chinn, A. Murdoch y G. Rubin

- 2009 "Epidemiology and management of infertility: a population-based study in UK primary care", en *Fam Pract*, England, vol. 26, pág. 269-74, DOI: [10.1093/fampra/cmp029](https://doi.org/10.1093/fampra/cmp029). (Citado en la pág. 116.)

Wong, WS y R Fielding

- 2011 "Prevalence and characteristics of chronic pain in the general population of Hong Kong", *J Pain*, 12, 2, pág. 236-45. (Citado en la pág. 121.)

Wu, C

- 2003 "Optimal Calibration Estimators in Survey Sampling", *Biometrika*, 90, 4, pág. 937-951. (Citado en la pág. 38.)
- 2004 "Some Algorithmic Aspects of the Empirical Likelihood Method in Survey Sampling", *Statistica Sinica*, 14, pág. 1057-1067. (Citado en la pág. 44.)

Wu, C y JNK Rao

- 2006 "Pseudo-empirical likelihood ratio confidence intervals for complex surveys", *Canadian Journal of Statistics*, 34, 3, pág. 359-375. (Citado en la pág. 85.)

Wu, Changbao y Randy R Sitter

- 2001 "A model-calibration approach to using complete auxiliary information from survey data", *J. Amer. Statist. Assoc.*, 96, 453, pág. 185-193, DOI: [10.1198/016214501750333054](https://doi.org/10.1198/016214501750333054). (Citado en las pág. 48, 69, 125.)

Wu, JS y CK Chu

- 1993 "Nonparametric function estimation and bandwidth selection for discontinuous regression functions", *Statistica Sinica*, 3, pág. 557-576. (Citado en la pág. 48.)

Yergens, Dean, Daniel Dutton y Scott Patten

- 2014 “An overview of the statistical methods reported by studies using the Canadian community health survey”, *BMC Medical Research Methodology*, 14, 1, pág. 15, <http://www.biomedcentral.com/1471-2288/14/15>. (Citado en la pág. 3.)

Yu, G., J. G. Sessions, Y. Fu y M. Wall

- 2015 “A multilevel cross-lagged structural equation analysis for reciprocal relationship between social capital and health”, *Soc Sci Med*, 142, pág. 1-8, DOI: [10.1016/j.socscimed.2015.08.004](https://doi.org/10.1016/j.socscimed.2015.08.004), <http://www.ncbi.nlm.nih.gov/pubmed/26277109>. (Citado en la pág. 3.)

Zardetto, D

- 2012 *R package EVER: Estimation of Variance by Efficient Replication*, pág. 89, <http://finzi.psych.upenn.edu/library/EVER/html/00Index.html>. (Citado en la pág. 92.)
- 2015 “ReGenesees: an Advanced R System for Calibration, Estimation and Sampling Error Assessment in Complex Sample Surveys”, *Journal of Official Statistics*, 31, 2, pág. 177-203, <http://dx.doi.org/10.1515/JOS-2015-0013>. (Citado en la pág. 95.)

Planteamiento general de la calibración de estimadores.

Sean:

- * $U = \{1, \dots, k, \dots, N\}$ la población de estudio constituida por N elementos distintos,
- * y_k la característica de interés asociada con el elemento k , y
- * s una muestra de U con valores asociados de y , $\{(k, y_k)\}$, seleccionada de acuerdo con un diseño muestral específico que asigna una probabilidad conocida $p(s)$, tal que $p(s) > 0$ para todo $s \in S$, conjunto de los posibles muestreos s , y $\sum_{s \in S} p(s) = 1$.

Supongamos que queremos estimar el total poblacional de la variable y , es decir, $T_y = \sum_{k \in U} y_k$, y que $x = (x_1, \dots, x_J)$ es un vector de J variables auxiliares, de forma que es perfectamente conocido para todos los elementos de U .

Entonces, si consideramos el estimador de Horvitz-Thompson para T_y , es decir, $\hat{T}_{yHT} = \sum_{k \in s} d_k y_k$, lo que pretendemos es modificar los pesos de diseño

$d_k = 1/\pi_k$, por otros pesos w_k , de forma que:

1. El estimador basado en dichos pesos proporciona estimaciones perfectas para x , es decir, $\sum_k w_k x_k = T_x = (T_{x_1}, \dots, T_{x_J})$, que conforman las ecuaciones de calibración.

2. Esos nuevos pesos calibrados w_k estén tan próximos como sea posible, respecto a una distancia dada, a los pesos originales d_k , es decir, $\min_{w_k} E \left\{ \sum_{k \in s} G(w_k, d_k) \right\}$, que constituirían las ecuaciones de equilibrio, cuya resolución se hace a través de los multiplicadores de Lagrange.