



ENCUESTA NACIONAL SOBRE LOS EFECTOS DEL COVID-19 EN EL BIENESTAR DE LOS HOGARES MEXICANOS (ENCOVID-19- OCTUBRE)

METODOLOGÍA

DISEÑO, SELECCIÓN Y CALIBRACIÓN DE LA MUESTRA

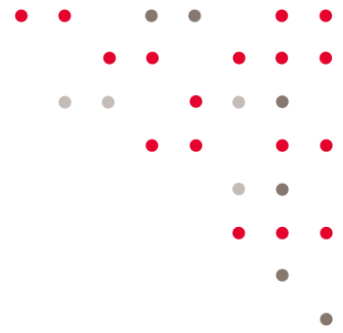
ANTECEDENTES

La Universidad Iberoamericana, A.C., a través del Instituto de Investigaciones para el Desarrollo con Equidad (EQUIDE), presenta los primeros resultados de la Encuesta de Seguimiento de los Efectos del COVID-19 en el Bienestar de los Hogares Mexicanos (ENCOVID-19). La **ENCOVID-19** fue diseñada por EQUIDE y QUANTOS Investigación Cuantitativa con el propósito de estudiar los impactos de la crisis sanitaria provocada por la pandemia del COVID-19 en el bienestar de los hogares mexicanos.

La **ENCOVID-19** busca proporcionar información para analizar los efectos de la contingencia sanitaria en el empleo, el ingreso, la salud mental y la nutrición de los hogares, ante el anuncio de las autoridades oficiales de que las encuestas presenciales sobre estos y otros temas serán suspendidas hasta nuevo aviso.¹ La **ENCOVID-19** proveerá de insumos fundamentales para entender la magnitud y severidad de la crisis del COVID-19 en la población mexicana, por lo que hacemos del conocimiento del público que dicha encuesta se llevará a cabo a principios de cada mes durante los siguientes 12 meses con el fin de captar los cambios observados en el bienestar de los hogares, que evolucionan rápidamente al paso de la crisis sanitaria.

El proyecto fue aprobado por el Comité de Ética en Investigación de la Universidad Iberoamericana con el registro: **CONBIOÉTICA-09—CEI-008-2016060**.

¹ <https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2020/especiales/INEGI-COVID.pdf>



DISEÑO MUESTRAL Y METODOLOGÍA

A continuación, se hacen explícitos elementos metodológicos del diseño de la ENCOVID-19:

- **Población objetivo:** Población de 18 años y más con telefonía móvil.
- **Marco muestral:** Números telefónicos móviles generados de manera aleatoria a partir de información del plan de marcación nacional. El marco muestral se mantiene actualizado inicio del levantamiento.
- **Técnica de recolección de datos:** Mediante entrevistas telefónicas a teléfonos móviles con cuestionario estructurado asistido por computadora (“*Computer Assisted Telephone Interviewing*” – **CATI**), aplicado por personal del EQUIDE capacitados exprofeso. En respuesta a la coyuntura actual, todos los encuestadores y supervisores del levantamiento telefónico trabajaron desde sus casas, sin riesgo de contagio y/o propagación del COVID19.
- **Diseño de muestreo:** Probabilístico unietápico de números telefónicos móviles. La selección de números telefónicos se realiza usando un diseño de muestreo aleatorio simple estratificado, implementado ‘*Random Digit Dialing*’ (**RDD**). Para la ejecución del RDD se utilizó software propio, específicamente el Paquete R RDDQuantosIC v1.5.
- **Extensión metodológica a hogares:** Interesa inferir a nivel hogar, por lo que se construyen las probabilidades de inclusión de los hogares a partir de información sobre el número de ocupantes con telefonía móvil en el hogar del informante.
- **Penetración de los servicios de telefonía:** De acuerdo con el IFT al 3er trimestre de 2019 se tiene una penetración de telefonía móvil del 96% en usuarios.
- **Dominios planeados de estimación:** Únicamente nacional.
- **Estratificación:** Se utiliza una estratificación por Entidad Federativa. Como resultado de estratificar a la población objetivo se obtiene un total de 32 estratos.
- **Distribución de muestra en estratos:** Se realizó de acuerdo con información de la Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares (ENDUTIH) 2018 de INEGI, según la distribución de usuarios de telefonía.



- **Conglomeración:** Al tratarse de un diseño de muestreo directo de elementos, es decir un diseño de muestreo unietápico, no hay conglomeración, por lo que teóricamente se tiene un diseño de muestreo (“*design effect*” – **DEFF**) menor a 1 en las estimaciones.
- **Márgenes de error:** El siguiente cuadro ofrece un cálculo de los márgenes de error asociados para diferentes niveles de la proporción estimada:

Proporción Estimada	Tamaño de Muestra Utilizado	Nivel de Confianza	Margen de Error Absoluto MAS bajo Supuesto de Normalidad
5%	1,600	95%	1.07%
10%	1,600	95%	1.47%
15%	1,600	95%	1.75%
20%	1,600	95%	1.96%
25%	1,600	95%	2.12%
30%	1,600	95%	2.25%
35%	1,600	95%	2.34%
40%	1,600	95%	2.40%
45%	1,600	95%	2.44%
50%	1,600	95%	2.45%
55%	1,600	95%	2.44%
60%	1,600	95%	2.40%
65%	1,600	95%	2.34%
70%	1,600	95%	2.25%
75%	1,600	95%	2.12%
80%	1,600	95%	1.96%
85%	1,600	95%	1.75%
90%	1,600	95%	1.47%
95%	1,600	95%	1.07%

- **Factores de expansión (pesos muestrales):** Los factores de expansión son el insumo principal para la generación de estimaciones. De manera inicial, los factores de expansión son el inverso de la probabilidad de inclusión. Estos se calculan considerando la distribución de muestra en los dominios y estratos.
- **Ajuste de factores de expansión:** Para evitar desviaciones en ciertas variables demográficas o socioeconómicas, los factores de expansión son ajustados por distribuciones geográficas, de edad y sexo de la población objetivo, proveniente de información auxiliar (censos, registros administrativos, estadísticas vitales, etcétera). En específico, se usó de información la Encuesta Inter Censal 2015 de INEGI y de la Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares



(ENDUTIH) 2018 de INEGI. Además, se utilizó información de nivel socioeconómico calculado de acuerdo a la definición 2018 de la AMAI.

- **Estimación y estimación de varianza:** Se utilizarán estimadores de expansión simple de Narain (1951) y Horvitz-Thompson (1952). Los factores de expansión correspondientes consideran una expansión al tamaño de la población acorde con su probabilidad de inclusión en muestra. Para los cálculos de estimación de varianza se propone el uso de expresiones directas, es decir, al no haber conglomeración no será necesario hacer uso del Principio de Conglomerados Últimos, e. g. Hansen, Hurwitz & Madow (1953) comúnmente utilizado por INEGI y otras agencias de estadística oficial.
- **Depurado y filtrado de números telefónicos:** Los números telefónicos generados de manera aleatoria incluyen números inactivos, contestadoras, etc., por lo que es necesario depurar estos de manera automática con el auto-marcador (proceso realizado por el CATI).
- **Filtros y controles específicos de levantamiento:** El diseño de muestreo es probabilístico, no obstante, implementó filtros de elegibilidad, y controles específicos de levantamiento (entrevistas sólo a mayores de edad, entrevistas exitosas para cada estrato, etc.). Asimismo, se controló la muestra para ciertos grupos de edad y para el sexo del entrevistado. El número específico para cada control se determinó de acuerdo con distribuciones e información de la población objetivo.
- **Levantamiento CATI durante la coyuntura COVID19:** Todos los encuestadores y supervisores trabajaron desde sus casas, sin riesgo de contagio y/o propagación del COVID19, con horarios regulares de trabajo y descansos, con procesos de supervisión. La marcación y el acopio de información es “en la nube”, y realizada por el EQUIDE.