



Universidad de San Carlos de Guatemala  
Escuela de Ciencias Físicas y Matemáticas  
Departamento de Matemática

4 **ELEMENTOS DE TEORÍA DE MUESTREO EN**  
5 **ENCUESTAS DE HOGARES**

6 **Sergio Alexander Alay Arellano**  
7 Asesorado por Cristian José Álvarez Bran

8 Guatemala, septiembre de 2024



10

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



11

ESCUELA DE CIENCIAS FÍSICAS Y MATEMÁTICAS

12

**ELEMENTOS DE TEORÍA DE MUESTREO EN  
ENCUESTAS DE HOGARES**

13

TRABAJO DE GRADUACIÓN

14

PRESENTADO A LA JEFATURA DEL

15

DEPARTAMENTO DE MATEMÁTICA

16

POR

17

**SERGIO ALEXANDER ALAY ARELLANO**

18

ASESORADO POR CRISTIAN JOSÉ ALVAREZ BRAN

19

AL CONFERÍRSELE EL TÍTULO DE

20

**LICENCIADO EN MATEMÁTICA APLICADA**

21

GUATEMALA, SEPTIEMBRE DE 2024





### CONSEJO DIRECTIVO INTERINO

Director	M.Sc. Jorge Marcelo Ixquiac Cabrera
Representante Docente	Arqta. Ana Verónica Carrera Vela
Representante Docente	M.A. Pedro Peláez Reyes
Representante de Egresados	Lic. Urías Amitaí Guzmán García
Representante de Estudiantes	Elvis Enrique Ramírez Mérida
Representante de Estudiantes	Oscar Eduardo García Orantes
Secretario	M.Sc. Freddy Estuardo Rodríguez Quezada

### TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

Director	Director
Examinador	Examinador 1
Examinador	Examinador 2
Examinador	Examinador 2
Secretario	Secretario



## AGRADECIMIENTOS





## DEDICATORIA



# ÍNDICE GENERAL

31	ÍNDICE DE FIGURAS	III
32	ÍNDICE DE TABLAS	V
33	LISTA DE SÍMBOLOS	VII
34	OBJETIVOS	IX
35	INTRODUCCIÓN	XI
36	<b>1. PRELIMINARES</b>	<b>1</b>
37	1.1. Teoría elemental de probabilidad . . . . .	1
38	1.1.1. ¿Qué es un espacio de probabilidad? . . . . .	1
39	1.1.2. Variables aleatorias . . . . .	1
40	1.1.2.1. Esperanza . . . . .	1
41	1.1.2.2. Varianza . . . . .	1
42	1.1.2.3. Función densidad . . . . .	1
43	1.1.2.4. Función distribución . . . . .	1
44	1.1.3. Distribuciones de probabilidad . . . . .	1
45	1.1.3.1. Binomial . . . . .	1
46	1.1.3.2. Normal . . . . .	1
47	1.1.3.3. Teorema del Límite Central . . . . .	1
48	1.1.4. Estimadores . . . . .	1
49	1.1.5. Intervalos de confianza . . . . .	1
50	<b>2. ¿Qué son las encuestas?</b>	<b>3</b>
51	2.1. ¿Qué elementos tiene una encuesta? . . . . .	3
52	2.2. Relevancia de las encuestas de hogares . . . . .	5
53	2.3. Errores muestrales y no muestrales . . . . .	6
54	2.4. Encuesta Multi-etápica . . . . .	10

55	2.4.1. Características de las Encuestas Multi-etápicas . . . . .	11
56	<b>3. Selección de la muestra</b>	<b>13</b>
57	3.1. Cálculo del tamaño de muestra . . . . .	13
58	3.1.1. ¿Cómo se define el tamaño de muestra de una encuesta a partir	
59	de la Teoría de probabilidad? . . . . .	14
60	3.2. Diseño muestral . . . . .	17
61	3.2.1. ¿Qué es un diseño muestral? . . . . .	17
62	3.2.2. Muestreo Aleatorio Simple . . . . .	18
63	3.2.3. Muestreo estratificado . . . . .	19
64	3.3. ¿Qué es el DEFF? . . . . .	21
65	3.3.1. Cálculo del DEFF en muestreo por conglomerados . . . . .	21
66	3.4. ¿Cómo se calcula el tamaño de muestra en encuestas de hogares? . . .	22
67	3.5. Algoritmos de selección . . . . .	23
68	3.5.1. Muestreo aleatorio simple . . . . .	23
69	3.5.2. Muestreo estratificado . . . . .	24
70	3.5.3. Muestreo por conglomerados . . . . .	24
71	3.5.4. Muestreo sistemático . . . . .	24
72	<b>4. Factores de expansión</b>	<b>25</b>
73	<b>CONCLUSIONES</b>	<b>27</b>
74	<b>RECOMENDACIONES</b>	<b>29</b>
75	<b>BIBLIOGRAFÍA</b>	<b>31</b>

## ÍNDICE DE FIGURAS



## ÍNDICE DE TABLAS





## LISTA DE SÍMBOLOS

Símbolo	Significado
$:=$	es definido por
$\cong$	es isomorfo a
$\Leftrightarrow$	si y sólo si
$\emptyset$	conjunto vacío
$E^c$	complemento de $E$
$\subsetneq$	estrictamente contenido
$E \setminus F$	diferencia entre $E$ y $F$
$E \Delta F$	diferencia simétrica entre $E$ y $F$
$\mathcal{P}(X)$	conjunto potencia de $X$
$\chi_E$	función característica de $E$
$E_n \uparrow$	$E_n$ es una sucesión creciente
$\mathfrak{L}$	$\sigma$ -álgebra de los conjuntos Lebesgue-medibles
$\mathcal{S}$	espacio muestral
$\mathfrak{A}$	$\sigma$ -álgebra de eventos
$(\mathcal{S}, \mathfrak{A}, P)$	espacio de probabilidad
$\mathcal{D}$	espacio de las funciones de prueba
$\mathcal{D}'$	espacio de las distribuciones
$\delta_0$	medida de Dirac, función $\delta$ de Dirac o $\delta$ -función
$\Phi^\times$	espacio antidual de $\Phi$
$\Phi \subset \mathcal{H} \subset \Phi^\times$	espacio de Hilbert equipado o tripleta de Gel'fand
$ \psi\rangle$	vector <i>ket</i>
$\langle\psi $	funcional <i>bra</i>
$\langle\varphi \psi\rangle$	<i>braket</i>



## OBJETIVOS

### General

Escriba el objetivo general.

### Específicos

Enumere los objetivos específicos.

1.

2.



# INTRODUCCIÓN



# 1. PRELIMINARES

## 1.1. Teoría elemental de probabilidad

### 1.1.1. ¿Qué es un espacio de probabilidad?

### 1.1.2. Variables aleatorias

#### 1.1.2.1. Esperanza

#### 1.1.2.2. Varianza

#### 1.1.2.3. Función densidad

#### 1.1.2.4. Función distribución

### 1.1.3. Distribuciones de probabilidad

#### 1.1.3.1. Binomial

#### 1.1.3.2. Normal

#### 1.1.3.3. Teorema del Límite Central

### 1.1.4. Estimadores

### 1.1.5. Intervalos de confianza

¿Cómo se definen? y en la práctica, ¿cómo se interpretan?





102

## 2. ¿Qué son las encuestas?

103 Consideraremos una encuesta como una herramienta de investigación que se  
104 utiliza para recolectar datos e información de un grupo específico de personas, cono-  
105 cido como muestra, con el objetivo de inferir características de la población general  
106 de donde se tomó la muestra. Una encuesta se caracteriza por utilizar instrumentos  
107 tales como cuestionarios o entrevistas, en los que se plantean preguntas estandari-  
108 zadas que permiten obtener respuestas comparables y cuantificables.

109 A lo largo de este capítulo se discutirán las generalidades, terminología e im-  
110 portancia de las encuestas.

### 111 2.1. ¿Qué elementos tiene una encuesta?

112 Con el propósito de asegurar la calidad y la efectividad de una encuesta, re-  
113 sulta fundamental considerar una serie de elementos clave que nos ayudarán para  
114 realizar su diseño, implementación y análisis. Estos elementos, además de garanti-  
115 zar la precisión de los datos recolectados, facilitan la interpretación y la aplicación  
116 práctica de los resultados y, principalmente, permiten generalizar a la población. A  
117 continuación se presentan los componentes que conforman una encuesta.

- 118 • **Unidad de observación.** Es el objeto sobre el que se realiza una medición.  
119 En el estudio de las poblaciones humanas, las unidades de observación suelen  
120 ser personas.
- 121 • **Población objetivo.** Es la colección completa de las unidades de observa-  
122 ción que deseamos estudiar. Usualmente, definir la población objetivo es una  
123 parte difícil del estudio<sup>1</sup> y la elección de esta afectará profundamente a las  
124 estadísticas resultantes.

---

<sup>1</sup>Por ejemplo, en un estudio de eficacia de un nuevo programa educativo ¿la población objetivo debe incluir solo a los estudiantes de un cierto grado escolar?, ¿deberíamos limitar la población objetivo a estudiantes de una región específica o, incluir estudiantes de todo el país? **intentá resumir esto.** Ya.

125 • **Muestra.** Subconjunto de la población que se utiliza para inferir o sacar  
126 conclusiones sobre toda la población.

127 • **Unidad de muestreo.** Es una unidad que puede ser seleccionada para for-  
128 mar parte de la muestra. Se puede dar la situación de que queramos estudiar  
129 personas, pero no dispongamos de una lista de todas las personas de la pobla-  
130 ción objetivo. En su lugar, los hogares sirven como unidades de muestreo y las  
131 unidades de observación son las personas que viven en los hogares.

132 • **Marco de muestreo.** Es una lista, un mapa u otra especificación de las unida-  
133 des de muestreo de la población a partir de las cuales puede seleccionarse una  
134 muestra. Es importante que este marco cuente con información que permita  
135 ubicar de manera pertinente<sup>2</sup> a cada una de las unidades que lo conforman.

136 Esto me gusta, pero no sé si aquí, dejemoslo por el momento y evaluemos si  
137 de repente hay un mejor lugar donde ponerlo Todos los marcos de muestreo  
138 presentan algún nivel de desactualización con respecto a la población objetivo.  
139 Por ejemplo, un marco de muestreo de áreas basado en cartografía, puede estar  
140 desactualizado debido a que es posible:

- 141 ○ entrevistar a la misma persona en varias ocasiones,
- 142 ○ nunca realizar la entrevista a una persona que no tiene un hogar fijo de  
143 residencia.

144 Esto también me gusta pero, de nuevo, no sé si aquí Es esencial saber que en las  
145 encuestas de personas, la población total suele ser menor que la población objetivo,  
146 es decir, no todas las peronas de la población objetivo están incluidas en el marco  
147 de muestreo, y varias personas no responderán a la encuesta.

148 Veamos cómo identificar algunos de estos elementos en dos de las encuestas  
149 que se han realizado en Guatemala.

150 **2.1 Ejemplo.** El Ministerio de Educación -MINEDUC- realizó el estudio “Situación  
151 nutricional de los escolares inscritos en los centros educativos públicos que son be-  
152 neficiarios del Programa de Alimentación Escolar, 2023” mediante la medición de  
153 peso y longitud/talla en una muestra estadísticamente significativa. En este estudio,  
154 las unidades de observación son los estudiantes matriculados en escuelas públicas,

---

<sup>2</sup>Para una encuesta de hogares con entrevistas *in situ*, las direcciones de los hogares de las personas a entrevistar. En el caso de una encuesta telefónica, los teléfonos de las personas a entrevistar.

155 ya que son los “objetos” sobre los cuales se realizaron las mediciones de peso y lon-  
156 gitud. La población objetivo del estudio son todos los estudiantes matriculados en  
157 escuelas públicas durante el año 2023. Esta es la colección completa de unidades de  
158 observación que se desea estudiar.

159 **2.2 Ejemplo.** La Encuesta de Evaluación de la Calidad de los Servicios Públicos  
160 Básicos -ENCASBA- 2019, brinda información acerca de los servicios públicos que  
161 proporciona el Organismo Ejecutivo, la Municipalidad de Guatemala y otras ins-  
162 tituciones públicas, la misma se sustenta en la evaluación que hizo la población  
163 informante del hogar, la cual podría ser el jefe o jefa del hogar o una persona de  
164 18 años o más de edad. La población objetivo está compuesta por las viviendas y  
165 los hogares del municipio de Guatemala. Dentro de este contexto, las unidades de  
166 muestreo son los hogares, ya que cada hogar puede ser seleccionado para formar  
167 parte de la muestra. La muestra es un subconjunto representativo de estos hogares  
168 seleccionados para inferir conclusiones sobre la percepción de la calidad de los servi-  
169 cios en toda la población objetivo del municipio. El marco de muestreo consiste en  
170 un listado detallado de todas las unidades primarias de muestreo (UPM) del muni-  
171 cipio, permitiendo identificar de manera precisa las unidades de muestreo elegibles  
172 para la encuesta.

173 Existen dos enfoques principales en el diseño de encuestas que se utilizan para  
174 recopilar datos sobre poblaciones:

- 175 • **Encuestas transversales.** Recopilan datos en un solo punto en el tiempo.  
176 Se utilizan para describir las características de una población en un momento  
177 específico y son útiles para realizar análisis de prevalencia o describir condicio-  
178 nes actuales. Como ejemplo de este enfoque podemos mencionar las *encuestas*  
179 *repetidas* que, además de ser una serie de encuesta aplicadas en diferentes mo-  
180 mentos, utilizan el mismo diseño metodológico. Para este tipo de encuestas, la  
181 selección de hogares se realiza de forma independiente para cada aplicación.
- 182 • **Encuestas longitudinales.** Siguen las mismas unidades de observación (co-  
183 mo hogares o individuos) a lo largo del tiempo. Esto permite analizar cambios y  
184 desarrollar comprensiones causales sobre cómo y por qué ocurren los cambios.

## 185 2.2. Relevancia de las encuestas de hogares

186 Las encuestas de hogares son utensilios esenciales en la investigación social,  
187 económica y demográfica. Proveen datos en diferentes tópicos:

- 188 1. *Salud y educación*. Proveen datos importantes sobre el acceso y calidad de  
189 servicios de salud y educación, lo que permite identificar desigualdades y áreas  
190 donde se requiere intervención. Esto es crucial para mejorar la cobertura y  
191 calidad de estos servicios (Incluir referencia: UNICEF - MICS).
- 192 2. *Investigación económica*. Datos sobre ingresos, gastos, empleo y otras variables  
193 económicas que son fundamentales para el análisis macro y microeconómico  
194 (Incluir referencia: OECD - Household Wealth).
- 195 3. *Calidad de Vida y Bienestar*. Recogen información sobre la calidad de vida y  
196 el bienestar subjetivo de las personas, incluyendo aspectos como la satisfacción  
197 con la vida, condiciones de vivienda, y acceso a servicios básicos. Ayudando  
198 así a construir un panorama integral del bienestar de la población (Incluir  
199 referencia: OECD - How's Life?).
- 200 4. *Desarrollo Sostenible*. En el contexto de los Objetivos de Desarrollo Sostenible  
201 -ODS-, las encuestas de hogares son una fuente de datos para monitorear el  
202 progreso hacia dichos objetivos (Incluir referencia: UN - SDG Indicators).

203 El diseño de la encuesta dependerá del objetivo de la medición, lo que se pretende al  
204 diseñar una encuesta de hogares es que esta sea un instrumento confiable, que brinde  
205 estimaciones exactas y precisas, ya que, de lo contrario, no se podrían monitorear  
206 de manera consistente las Políticas Públicas y los indicadores de interés.

## 207 2.3. Errores muestrales y no muestrales

208 En este apartado se trata de describir el paradigma de los errores que se cometen  
209 en una encuesta y cómo medirlos de manera acertada y acotarlos sobre la base del  
210 *principio de representatividad*. Para esto empezamos definiendo dicho principio.

211 *Principio de representatividad en encuestas*. Afirmar que cada elemento incluido  
212 en una muestra se representa a sí mismo y a un grupo de unidades que no pertenecen  
213 a la muestra seleccionada, cuyas características son cercanas a las del elemento  
214 incluido en la muestra

215 Este principio es crucial para garantizar que los resultados de la encuesta sean  
216 generalizables y reflejen con precisión las opiniones, comportamientos o caracterís-  
217 ticas de toda la población objetivo. Existen dos fuentes principales de error cuando  
218 se realiza una encuesta:

- 219 1. **Error muestral.** Este tipo de error ocurre porque no se incluyó a todas  
220 las personas de la población y se seleccionó una muestra. La magnitud del  
221 error muestral depende del tamaño de la muestra y del método de muestreo  
222 utilizado. En general, un tamaño de muestra mayor tiende a reducir el error  
223 muestral.
- 224 2. **Error no muestral.** Este tipo de error abarca todos los demás errores que  
225 pueden ocurrir en una encuesta que no están relacionados con el hecho de que  
226 se ha utilizado una muestra en lugar de un censo. Estos errores pueden surgir  
227 en cualquier etapa del proceso de la encuesta, desde el diseño del cuestionario  
228 hasta la recolección y el análisis de los datos. El error no muestral puede  
229 ser sistemático o aleatorio. Entre los diferentes tipos de errores no muestrales  
230 están:
- 231 • *Error de cobertura.* Ocurre cuando algunos miembros de la población  
232 objetivo no tienen la oportunidad de ser seleccionados en la muestra.
  - 233 • *Error de respuesta.* Surge cuando los encuestados no proporcionan res-  
234 puestas precisas, ya sea intencionalmente o por confusión; malentendidos  
235 del cuestionario o respuestas socialmente deseables.
  - 236 • *Error de no respuesta.* Se produce cuando ciertas personas seleccionadas  
237 para la muestra no responden a la encuesta, y estas personas pueden  
238 tener características u opiniones diferentes a las de quienes sí responden.
  - 239 • *Errores de procesamiento.* Incluyen errores en la entrada de datos, codi-  
240 ficación o análisis de los datos de la encuesta.

241 Dentro de los errores no muestrales existe un tipo de error denominado **sesgo** que  
242 afecta la validez de los resultados de la encuesta; es una tendencia o inclinación  
243 que distorsiona la verdad de la representatividad de los datos, llevándolos a ser no  
244 objetivos o inexactos. Los sesgos pueden ocurrir en cualquier etapa del proceso de  
245 la encuesta: desde el diseño hasta la recolección y análisis de los datos.

246 Según [11] las dos fuentes más importantes de sesgo se resumen de la siguiente  
247 manera:

- 248 1. **Sesgo de selección.** Ocurre cuando parte de la población objetivo no es-  
249 tá en el marco de muestreo o cuando el marco está incompleto y presenta  
250 deficiencias. En [1] se establece que este tipo de sesgo se presenta si:

- La selección de la muestra depende de cierta característica asociada a las propiedades de interés. Por ejemplo, la frecuencia con la que los adolescentes hablan con los padres acerca del SIDA.
- La muestra se realiza mediante elección deliberada o mediante un juicio subjetivo. Por ejemplo, si el parámetro de interés es la cantidad promedio de gastos en compras en un centro comercial y el encuestador elige a las personas que salen con muchos paquetes, entonces la información estaría sesgada puesto que no está reflejado el comportamiento verdadero de las compras.
- Existen errores en la especificación de la población objetivo. Por ejemplo, en encuestas electorales, cuando la población objetivo contiene a personas que no están registradas como votantes ante la organización electoral de su país.
- Existe sustitución deliberada de unidades no disponibles en la muestra. Si, por alguna razón, no fue posible obtener la medición y consecuente observación de la característica de interés para algún individuo en la población, la sustitución de este elemento debe hacerse bajo estrictos procedimientos estadísticos y no debe ser subjetiva en ningún modo.
- Existe ausencia de respuesta. Este fenómeno puede causar distorsión de los resultados cuando los que no responden a la encuesta difieren críticamente de los que sí respondieron.
- La muestra está compuesta por respondientes voluntarios. Los foros radiales, las encuestas de televisión y los estudios de portales de internet no proporcionan información confiable.

**2. Sesgo de medición.** Se da cuando el instrumento con el que se realiza la medición tiene una tendencia a diferir del valor verdadero que se desea averiguar. Éste sesgo debe ser considerado y minimizado en la etapa de diseño de la encuesta. Nótese que ningún análisis estadístico puede revelar que una pesa añadió a cada persona dos kilos de más en un estudio de salud. De igual manera, en [1] se encuentran algunas de las situaciones en las que se presenta el sesgo de medición:

- Cuando el respondiente miente. Esta situación se presenta a menudo en encuestas que se pregunta acerca del ingreso salarial, alcoholismo y drogadicción, nivel socioeconómico e incluso edad.

- 285 • Difícil comprensión de las preguntas. Por ejemplo: ¿No cree que no esté  
286 en buen momento para invertir? La doble negación en la pregunta es muy  
287 confusa para el respondiente.
- 288 • Las personas tienden a olvidar. Es bien sabido que las malas experiencias  
289 suelen ser olvidadas; esta situación debe acotarse si se está trabajando  
290 en una encuesta de criminalidad.
- 291 • Distintas respuestas a distintos entrevistadores. En algunas regiones es  
292 muy probable que la raza, edad o género del encuestador afecte directa-  
293 mente la respuesta del entrevistado.
- 294 • Leer mal las preguntas o polemizar con el respondiente. El encuestador  
295 puede influir notablemente en las respuestas. Por lo anterior, es muy im-  
296 portante que el proceso de entrenamiento del entrevistador sea riguroso  
297 y completo.

298 El siguiente ejemplo ilustra cómo es que los sesgos afectan a una encuesta.

299 **2.3 Ejemplo.** Durante las elecciones generales de 2015 en Guatemala, se llevaron  
300 a cabo varias encuestas preelectorales para predecir los resultados de la primera  
301 vuelta presidencial. Sin embargo, los resultados de las encuestas no coincidieron con  
302 los resultados finales, generando críticas y cuestionamientos sobre la metodología  
303 empleada. Entre los problemas que se identificaron, tenemos:

#### 304 1. Muestra no representativa

- 305 • *Subrepresentación<sup>3</sup> de votantes rurales:* Era recurrente que las encuestas  
306 no incluían adecuadamente a los votantes en áreas rurales, que represen-  
307 tan una parte significativa del electorado en Guatemala. Esto llevó a un  
308 sesgo en los resultados hacia las preferencias de los votantes urbanos.
- 309 • *Dificultades logísticas:* La geografía y el acceso limitado a algunas re-  
310 giones hicieron que las encuestas no pudieran captar completamente las  
311 opiniones de todas las áreas del país.

#### 312 2. Tasa de respuesta baja

- 313 • *Desconfianza en las encuestas:* Muchos ciudadanos mostraron descon-  
314 fianza hacia las encuestas, lo que resultó en tasas de respuesta bajas y  
315 potencialmente sesgadas.

---

<sup>3</sup>explicar qué significa esto

- *Acceso a tecnología:* En algunas encuestas que utilizaron métodos en línea o telefónicos, hubo subrepresentación de personas sin acceso a tecnología, como teléfonos o internet.

### 3. Errores en la proyección de participación

*Participación inesperada:* Las encuestas no lograron capturar cambios en la participación electoral, como el aumento en el número de votantes jóvenes o de primer voto, que influyeron en los resultados finales.

### 4. Formulación de preguntas

*Influencias en las respuestas:* Las preguntas utilizadas en algunas encuestas podrían haber sido percibidas como sesgadas o influyentes, lo que podría afectar la sinceridad de las respuestas.

## 2.4. Encuesta Multi-etápica

Las encuestas multi-etápicas son un tipo de diseño muestral utilizado en la investigación estadística y de encuestas. En este enfoque, la selección de la muestra se realiza en varias etapas, en lugar de seleccionar directamente a los individuos de la población objetivo. Este método es particularmente útil cuando se trata de poblaciones grandes y dispersas, donde sería costoso o logísticamente difícil realizar un muestreo directo. En general, en América Latina son muy comunes los diseños de selección en dos etapas:

- **Primera etapa.** Se seleccionan unidades grandes, llamadas Unidades Primarias de Muestreo (UPM), que pueden ser regiones geográficas, bloques o conglomerados.
- **Segunda etapa.** Dentro de cada UPM se seleccionan unidades más pequeñas como viviendas u hogares.

También es posible encontrar en algunos países diseños divididos en más de dos etapas. Por ejemplo, en una primera etapa se seleccionan municipios, en una segunda etapa se seleccionan UPM dentro de los municipios seleccionados y, en la tercera, se selecciona una muestra de hogares en aquellas UPM seleccionadas en la segunda etapa pertenecientes a los municipios seleccionados en la primera etapa de muestreo.



### 345 2.4.1. Características de las Encuestas Multi-etápicas

346 Según [10], el muestreo multi-etápico presenta dos características fundamenta-  
347 les que lo hacen estadísticamente robusto y eficiente en la planificación logística del  
348 proceso de recopilación de datos

- 349 • **La independencia**, que implica que no hay ninguna correlación en el diseño  
350 de muestreo de las UPM. Esto quiere decir que en cada UPM se puede ejecutar  
351 con independencia cualquier estrategia de muestreo que se considere apropiada  
352 para seleccionar la submuestra de hogares.
- 353 • **La varianza**, que implica que sin importar qué diseño de muestreo se haya  
354 ejecutado en la primera etapa para seleccionar las UPM, la segunda etapa de  
355 selección podrá ejecutarse de manera independiente a la primera. Es decir, el  
356 submuestreo de los hogares es independiente del muestreo de las UPM.



### 3. Selección de la muestra

358 El cálculo del tamaño de muestra es un elemento fundamental en la planifica-  
359 ción y ejecución de encuestas, ya que determina la precisión y la representatividad  
360 de los resultados obtenidos. En este capítulo se abordará el proceso de determinar  
361 el tamaño de muestra óptimo basándose en la Teoría de Probabilidad. Este enfo-  
362 que permite asegurar que la muestra seleccionada sea suficientemente grande para  
363 reflejar las características de la población, pero también lo suficientemente pequeña  
364 para ser práctica y costo-efectiva.

#### 3.1. Cálculo del tamaño de muestra

366 Vamos a tratar de abordar el problema de *determinar el valor de alguna can-*  
367 *tidad o indicador de la población objetivo (dichos valores pueden ser: ingreso total,*  
368 *cantidad total de personas, proporción de personas con cierta característica, índice*  
369 *total de alguna característica de las personas, etc).* El camino seguro pero exhaustivo  
370 sería el de recabar información de toda la población para poder conocer el indicador  
371 deseado, pero, desde ya, esto tiene demasiadas desventajas: económicas, logísticas,  
372 tiempo, errores de cobertura, calidad de los datos. Por lo tanto, lo ideal es realizar  
373 la recabación de información en solamente un pequeño porcentaje de la población  
374 para lograr determinar el valor del indicador en cuestión.

375 Pretender usar una cantidad limitada de información para determinar con pre-  
376 cisión un valor exacto a nivel poblacional es poco realista. Pero es razonable utilizar  
377 esa información para obtener una idea general de la situación. Por esto, en lugar de  
378 buscar un valor exacto, nos enfocamos en determinar un rango que contenga el valor  
379 del indicador, algo como “La cantidad total de personas en el país está entre 15 y 20  
380 millones”. Idealmente, quisiéramos que este rango sea lo más estrecho posible para  
381 tener una estimación más concreta del valor real del indicador.

382 Sin embargo, para hacerlo correctamente debemos superar obstáculos clave  
383 como asegurar que la cantidad de información disponible sea suficiente, que esta

información refleje fielmente la realidad de la población, y encontrar una manera de evaluar la calidad de la estrategia empleada. Por esta razón utilizamos la Teoría de la Probabilidad como marco metodológico, ya que nos proporciona las herramientas necesarias para realizar este ejercicio de manera fundamentada, considerando toda la lógica mencionada.

### 3.1.1. ¿Cómo se define el tamaño de muestra de una encuesta a partir de la Teoría de probabilidad?

Sea  $s$  una muestra de  $n \in \mathbb{N}$  individuos que pertenecen a una población  $X = \{x_1, x_2, \dots, x_N\}$ , y sea  $\mathcal{S}$  la colección de todas esas posibles muestras  $s$ . Cada individuo  $x_i$  tiene asociado un número real  $y_i$  que captura alguna cantidad del parámetro poblacional  $\theta$ , por ejemplo, el ingreso del individuo que contribuye al ingreso promedio de la población

$$\frac{1}{N} \sum_{i=1}^N y_i. \quad (3.1)$$

Consideremos una función  $\hat{\theta}: \mathcal{S} \rightarrow \mathbb{R}$  que asigna a cada posible muestra  $s \in \mathcal{S}$  un valor real, que representará la información pertinente para el objetivo de estimar al parámetro poblacional y que se puede obtener de la medición de la muestra seleccionada. Por ejemplo, si el objetivo es ubicar el ingreso promedio de  $X$ , una función natural a considerar como punto de partida es el ingreso promedio de la muestra:

$$\hat{\theta}(s) = \frac{1}{n} \sum_{x_i \in s} y_i \quad (3.2)$$

donde  $y_i$  corresponde al ingreso del individuo  $x_i$ .

La función es “natural” en el sentido de que emula en la muestra lo que se desea saber de la población, pero a parte de ello no tiene por qué ser necesariamente una función útil para ubicar a (3.1), puede que la muestra  $s$  seleccionada tenga características muy singulares respecto al resto de la población y  $\hat{\theta}(s)$  no esté ni cerca del valor poblacional. Sin embargo, si consideramos esta función en el contexto más general de todas las posibles muestras  $s$  que se pueden obtener, al promediar

409 todos los resultados de calcular  $\hat{\theta}(s)$  se llega a que

$$\begin{aligned}
 \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \hat{\theta}(s) &= \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \frac{1}{n} \sum_{x_i \in s} y_i \\
 &= \frac{1}{|\mathcal{S}|} \sum_{i=1}^N \frac{y_i}{n} \cdot |\{s \in \mathcal{S} : x_i \in s\}| \\
 &= \frac{1}{\binom{N}{n}} \sum_{i=1}^N \frac{y_i}{n} \binom{N-1}{n-1} \\
 &= \frac{1}{N} \sum_{i=1}^N y_i.
 \end{aligned} \tag{3.3}$$

410 Es decir, el promedio de los  $\hat{\theta}(s)$  es exactamente igual al parámetro deseado. A  
 411 primera vista esto quizás pueda parecer, como mucho, un resultado bonito e intere-  
 412 sante pero que carece de utilidad práctica. Después de todo, disponer de  $\hat{\theta}(s)$  para  
 413 toda posible muestra es equivalente a tener la información de toda la población,  
 414 así que hemos vuelto al punto de partida. Sin embargo, intuitivamente el promedio  
 415 de ciertas cantidades se interpreta como esa medida “resumen” o “punto medio” del  
 416 contexto general de esas cantidades y esta noción se potencia en el contexto de la  
 417 Teoría de Probabilidad. Si  $\mathcal{S}$  es un espacio de probabilidad, es decir, cada  $s \in \mathcal{S}$  tie-  
 418 ne asignada una probabilidad  $p(s)$  de ser la muestra seleccionada, entonces  $\hat{\theta}$  es una  
 419 variable aleatoria. Si además, cada  $s$  tiene la misma probabilidad de ser la muestra

$$p(s) = \frac{1}{|\mathcal{S}|}$$

420 entonces el cálculo que acabamos de hacer nos dice que  $f$  es un estimador insesgado  
 421 del parámetro poblacional, es decir

$$\begin{aligned}
 E[\hat{\theta}] &= \sum_{s \in \mathcal{S}} \frac{1}{|\mathcal{S}|} f(s) \\
 &= \frac{1}{N} \sum_{i=1}^N y_i \\
 &= \theta.
 \end{aligned}$$

422 Lo interesante de esto es que es posible ubicar la esperanza de una variable aleatoria  
 423 en cierto rango con cierta “confianza” asociada a esta estructura probabilística que  
 424 existe en  $\mathcal{S}$ , gracias primordialmente al famoso Teorema del Límite Central. En

425 resumen, por medio de un estimador insesgado  $\hat{\theta}$  del parámetro poblacional  $\theta$ , existe  
 426 una buena forma de ubicar a  $\theta$  y todas las tareas alrededor de la muestra estarán  
 427 íntimamente ligadas a esto.

428 En las encuestas se considera un tamaño fijo de elementos de una población  
 429 para pertenecer a una muestra  $s$  sobre la cual se recopilará información. Una de las  
 430 primeras tareas al diseñar una encuesta consiste en definir ese tamaño de elementos  
 431 de la población, esto se hace al **ubicar el parámetro de interés** definido por el  
 432 objetivo principal de la encuesta.

433 Ubicar el parámetro  $\theta$  consiste en construir un intervalo de confianza para  $\theta$   
 434 con un nivel de significancia  $\alpha$  y una amplitud  $A$ . Tanto  $\alpha$  como  $A$  son cantidades  
 435 que pueden depender del criterio de las personas interesadas en los resultados de  
 436 la encuesta puesto que son quienes tienen, o deberían tener, una noción adecuada  
 437 de las dimensiones de las cantidades recopiladas, además de las aplicaciones que se  
 438 desprenden de los resultados de la encuesta, por ejemplo: la elaboración de Políticas  
 439 Públicas o los análisis socioeconómicos (Manual de Encuestas Sobre Hogares).

440 Por el Teorema del Límite Central tenemos que para un tamaño de muestra  
 441  $|s| = n$  suficientemente grande en una población grande y un estimador insesgado  $\hat{\theta}$   
 442 de  $\theta$  se cumple que

$$P\left(\left|\hat{\theta} - \theta\right| \leq c\sqrt{V(\hat{\theta})}\right) \approx \phi(c), \quad c \in \mathbb{R}$$

443 donde  $\phi$  es la función de distribución de una variable aleatoria normal estándar (en  
 444 otras palabras,  $\phi(c)$  es un valor conocido). Esto implica que el intervalo

$$I = \left[\hat{\theta} - c\sqrt{V(\hat{\theta})}, \hat{\theta} + c\sqrt{V(\hat{\theta})}\right] \quad (3.4)$$

445 contiene al parámetro  $\theta$  con probabilidad  $\phi(c)$ . Como podemos ver en (3.4), la am-  
 446 plitud  $A$  de  $I$  depende de  $c$ , que está ligado al nivel de significancia que usualmente  
 447 se establece en un 95 %, y de la varianza  $V(\hat{\theta})$  del estimador.

448 Supongamos que la varianza del estimador es una función decreciente del tama-  
 449 ño de muestra  $n$ , es decir  $V(\hat{\theta}) = h(n)$  con  $n: \mathbb{N} \rightarrow \mathbb{R}$  tal que  $a < b \implies h(a) > h(b)$ .

450 Si se busca que la amplitud de  $I$  no supere un umbral  $A_0$ , entonces

$$\begin{aligned} 2c\sqrt{V(\hat{\theta})} &\leq A_0 \iff \\ h(n) &\leq \left(\frac{A_0}{2c}\right)^2 \iff \\ n &\geq h^{-1}\left(\left(\frac{A_0}{2c}\right)^2\right). \end{aligned}$$

451 Es decir, podemos hallar el tamaño “óptimo” de muestra tomando el  $n$  más pequeño  
452 que satisface la desigualdad anterior.

453 La forma que toma  $V(\hat{\theta})$  dependerá de la estructura probabilística de  $\mathcal{S}$ , así  
454 como del estimador  $\hat{\theta}$ .

455 Anteriormente vimos que un diseño de muestreo consiste en una medida de  
456 probabilidad definida en  $\mathcal{S}$ , que consiste en asignar a cada  $s \in \mathcal{S}$  un número  $p(s) \in$   
457  $(0, 1)$  que captura “qué tan probable” es que  $s$  sea *la* muestra.

## 458 3.2. Diseño muestral

459 En el ámbito de las encuestas de hogares, el diseño muestral ocupa un lugar  
460 central como fundamento de los procesos inferenciales que permiten obtener infor-  
461 mación precisa y confiable sobre la población de un país. La premisa básica de este  
462 enfoque radica en que una muestra representativa puede reflejar las características  
463 más importantes de una población extensa. Esta idea, aunque ambiciosa, se respal-  
464 da en la implementación de los procedimientos robustos y metodologías estadísticas  
465 rigurosas que garantizan estimaciones válidas y exactas, incluso a partir del estudio  
466 de una fracción relativamente pequeña de la población objetivo.

467 El diseño muestral, por tanto, no solo facilita el levantamiento de datos con  
468 eficiencia operativa y presupuestaria, sino que también constituye la base para ge-  
469 nerar indicadores sociales y económicos fundamentales. Estos indicadores permiten  
470 evaluar Políticas Públicas, monitorear el desarrollo y guiar la toma de decisiones  
471 informadas para mejorar las condiciones de vida de la sociedad.

### 472 3.2.1. ¿Qué es un diseño muestral?

473 Un diseño muestral es un conjunto de procedimientos y estrategias que definen  
474 la forma en que se seleccionará una muestra a partir de una población objetivo. En  
475 términos prácticos, es el proceso que garantiza que cada unidad de la población tenga

una probabilidad conocida y no nula de ser incluida en la muestra. Esto permite que las estimaciones derivadas sean representativas y precisas, evitando sesgos y proporcionando bases sólidas para realizar inferencias estadísticas.

El diseño muestral se construye a partir de principios fundamentales como la aleatorización y la inclusión:

- **Aleatorización:** Las unidades que componen la muestra se seleccionan de forma probabilística, asegurando que cualquier combinación plausible de hogares o individuos pueda ser parte de la muestra.
- **Inclusión:** Todas las unidades de la población tienen una probabilidad no nula de ser seleccionadas, garantizando que ninguna unidad quede excluida del proceso.

### 3.2.2. Muestreo Aleatorio Simple

El diseño muestral más sencillo de todos, sobre el cual toda la teoría parte, es el Muestreo Aleatorio Simple. La definición consiste en considerar cada una de las  $s \in \mathcal{S}$  como igual de probables, es decir:

$$p(s) = \frac{1}{|\mathcal{S}|}. \quad (3.5)$$

Bajo este diseño, el estimador natural de un promedio poblacional resulta tener varianza

$$V_{\text{MAS}}(\hat{\theta}) = \frac{S^2}{n} \left(1 - \frac{n}{N}\right)$$

donde  $S^2$  y  $N$  son constantes de la población, y  $n$  es el tamaño de la muestra. Se puede verificar que esta es una función decreciente en  $n$ , así que por lo descrito anteriormente existe una forma sistemática en la cual uno define el tamaño de muestra para una encuesta bajo este diseño:

$$n = \frac{S^2}{\left(\frac{A_0}{2}\right)^2 \cdot \frac{1}{c^2} + \frac{S^2}{N}} \quad (3.6)$$

Al hacer  $A_0/2 = e$ ,  $c = z_{\alpha/2}$  y suponiendo que  $N$  es mucho más grande que  $S$  se llega a la fórmula que se encuentra en los libros de texto:

$$n = \left(\frac{z_{\alpha/2} S}{e}\right)^2.$$



### 499 3.2.3. Muestreo estratificado

500 El muestreo estratificado es una técnica de muestreo que mejora la precisión de  
501 las estimaciones al dividir la población en subgrupos homogéneos llamados **estratos**.  
502 Estos estratos, que a menudo corresponden a grandes unidades geográficas, permiten  
503 garantizar una representación adecuada de cada grupo en la muestra, facilitando así  
504 inferencias más precisas sobre subpoblaciones específicas de interés.

505 Dividimos la población de  $N$  en  $H$  estratos, con  $N_h$  unidades de muestreo en el  
506 estrato  $h$ . Para que el muestreo estratificado funcione, debemos conocer los valores  
507 de  $N_1, N_2, \dots, N_H$  y debe cumplirse que

$$N_1 + N_2 + \dots + N_H = N$$

508 donde  $N$  es el número total de unidades en toda la población. Cada estrato es tratado  
509 como una subpoblación independiente, y se selecciona una muestra de tamaño  $n_h$   
510 dentro de cada estrato, asegurando que

$$n_1 + n_2 + \dots + n_H = n.$$

511 La asignación de la muestra a los estratos puede hacerse de distintas maneras, de-  
512 pendiendo de los objetivos de la encuesta y de la variabilidad dentro de cada estrato.

- 513 1. **Asignación simple o uniforme.** Cada estrato recibe el mismo tamaño de  
514 muestra sin importar su tamaño en la población:

$$n_h = \frac{n}{H}$$

515 Este método es poco eficiente si los estratos tienen tamaños muy distintos o  
516 varianzas diferentes.

- 517 2. **Asignación proporcional.** Se asigna el tamaño de muestra a cada estrato  
518 de acuerdo con su peso en la población

$$n_h = n \cdot \frac{N_h}{N}.$$

519 Este método es útil cuando los estratos tienen tamaños muy diferentes pero  
520 similar variabilidad.

- 521 3. **Asignación óptima (Neyman).** Si la variabilidad dentro de los estratos es  
522 diferente, se usa la asignación óptima de Neyman, que minimiza la varianza

del estimador global

$$n_h = n \cdot \frac{N_h S_h}{\sum_{h=1}^H N_h S_h},$$

donde  $S_h$  es la desviación estándar de la variable de interés dentro del estrato  $h$ . Esta asignación es más eficiente porque asigna más muestra a los estratos con mayor variabilidad, mejorando la precisión del estimador.

La estimación del promedio poblacional  $\bar{Y}$  en un diseño estratificado se calcula como

$$\hat{Y}_s = \sum_{h=1}^H W_h \hat{Y}_h,$$

donde:

- $W_h = N_h/N$  es el peso del estrato en la población,
- $\bar{Y}_h$  es la media muestral del estrato  $h$ .

La varianza del estimador es:

$$V(\hat{Y}_s) = \sum_{h=1}^H W_h^2 \frac{S_h^2}{n_h} \left(1 - \frac{n_h}{N_h}\right).$$

Si la asignación de la muestra se hace con la regla de Neyman, se minimiza esta varianza, logrando una estimación más eficiente.

El tamaño de muestra total  $n$  para un muestreo estratificado se puede determinar a partir de la varianza deseada y del margen de error permitido en la estimación. Si el objetivo es estimar una media con error estándar máximo  $\epsilon$ , el tamaño de muestra se obtiene resolviendo la ecuación

$$n = \frac{\sum_{h=1}^H W_h^2 S_h^2}{\epsilon^2 / z^2}. \quad (3.7)$$

Si se estima una proporción poblacional  $P$ , el tamaño de muestra se puede aproximar por

$$n = \frac{\sum_{h=1}^H W_h^2 P_h (1 - P_h)}{\epsilon^2 / z^2}. \quad (3.8)$$

Donde:

- $z$  es el valor crítico de la distribución normal estándar según el nivel de confianza deseado (por ejemplo, 1.96 para 95 %),
- $\epsilon$  es el margen de error máximo permitido.

### 544 3.3. ¿Qué es el DEFF?

545 El efecto de diseño (DEFF, por sus siglas en inglés “Design Effect”) es una me-  
546 dida que cuantifica cuánto más grande debe ser una muestra obtenida con un diseño  
547 muestral complejo (como el estratificado, por conglomerados, o una combinación  
548 de ambos) en comparación con una muestra obtenida mediante muestreo aleatorio  
549 simple, para lograr la misma precisión en las estimaciones.

550 Se define como

$$\text{DEFF} = \frac{V(\hat{\theta}_{\text{complejo}})}{V(\hat{\theta}_{\text{MAS}})}, \quad (3.9)$$

551 donde:

- 552 •  $V(\hat{\theta}_{\text{complejo}})$  es la varianza del estimador en el diseño muestral utilizado (estra-  
553 tificado, por conglomerados, etc.).
- 554 •  $V(\hat{\theta}_{\text{MAS}})$  es la varianza del estimador si se hubiera usado un muestreo aleatorio  
555 simple.

556 Un  $\text{DEFF} > 1$  indica que el diseño muestral utilizado requiere una muestr mayor  
557 para alcanzar la misma precisión que el MAS, muestras que un  $\text{DEFF} < 1$  indica  
558 que el diseño muestral es más eficiente que el MAS.

#### 559 3.3.1. Cálculo del DEFF en muestreo por conglomerados

560 En el caso del muestreo por conglomerados, el DEFF tiende a ser mayor que 1  
561 porque los individuos dentro de un conglomerado tienden a ser más similares entre  
562 sí que si fueran seleccionados aleatoriamente. En este caso, el DEFF se aproxima  
563 como:

$$\text{DEFF} \approx 1 + (b - 1)\rho \quad (3.10)$$

564 donde:

- 565 •  $\rho$  es el coeficiente de correlación intraclase, que mide la similitud dentro de los  
566 conglomerados.
- 567 •  $b$  es el tamaño promedio de los conglomerados.

568 Si la correlación intraclase  $\rho$  es alta, significa que los individuos dentro de cada  
569 conglomerado son muy similares, lo que aumenta la varianza y hace que el DEFF sea

mayor que 1, reduciendo la eficiencia del diseño. Cuando se usa un diseño muestral complejo, el tamaño de muestra necesario se ajusta multiplicándolo por el DEFF,

$$n_{\text{ajustado}} = n_{\text{MAS}} \times \text{DEFF}.$$

Donde:

- $n_{\text{MAS}}$  es el tamaño de muestra que se habría usado con muestreo aleatorio simple.

- $n_{\text{ajustado}}$  es el tamaño de muestra corregido por el diseño muestral.

Si el  $\text{DEFF} > 1$ , se necesita un tamaño de muestra mayor para mantener la precisión deseada en las estimaciones.

### 3.4. ¿Cómo se calcula el tamaño de muestra en encuestas de hogares?

El cálculo del tamaño de muestra para la mayoría de encuestas de hogares inicia con el planteamiento de un diseño muestral adecuado y este comienza con la definición de los objetivos de la encuesta. Esto permite delimitar aspectos como: la población objetivo, los dominios de estudio, la variable de diseño y los estratos. Una vez fijados estos elementos se procede al cálculo del tamaño de la muestra necesario a tener, es decir, unidades informantes sobre las que hay que recolectar información. En encuestas de gran alcance, en las que se pretende obtener información en varias etapas: primero se seleccionan sectores cartográficos (Unidades Primarias de Muestreo) y luego hogares dentro de estos sectores seleccionados (Unidades Secundarias de Muestreo).

El tamaño de muestra está íntimamente asociado a la precisión que se desee obtener en la estimación del parámetro de interés (variable de diseño). Esta precisión se define tomando en consideración el contexto de la información a recopilar y las capacidades logísticas y presupuestarias disponibles. El proceso del cálculo de este tamaño debe ser realizado en cada dominio de estudio a considerar. Finalmente, una vez calculado el tamaño de muestra, se define una estrategia de afijación de la muestra, es decir, en cada dominio de estudio se distribuye el tamaño de muestra calculado en los estratos que conforman el dominio.

En particular, cuando se trabaja en R, existe un paquete muy útil llamado *samplesize4surveys* el cual contiene una serie de funciones que consideran el tipo

600 de la variable de diseño y las etapas del diseño muestral. Por lo tanto, resulta muy  
601 sencillo realizar una rutina en R para determinar el tamaño muestral en cada dominio  
602 de estudio para la encuesta en cuestión. Entonces, lo principal es establecer

## 603 3.5. Algoritmos de selección

604 Los algoritmos de selección son la base operativa de los diseños muestrales, per-  
605 mitiendo la implementación sistemática de distintos métodos para elegir unidades  
606 de análisis dentro de una población. Su función principal es garantizar que la selec-  
607 ción de la muestra se realice de manera precisa, objetiva y conforme a los principios  
608 probabilísticos que rigen la inferencia estadística.

609 Existen diversos algoritmos de selección, cada uno asociado a un diseño mues-  
610 tral específico. Los más comunes son: el de *muestreo aleatorio simple*, el *muestreo*  
611 *estratificado*, el *muestreo por conglomerados*, y el *muestreo sistempático*. En cada  
612 caso, los algoritmos determinan la forma en que se eligen las unidades, asegurando  
613 que la muestra cumpla con los requisitos de aleatorización y representatividad.

614 El diseño de un algoritmo de selección depende de varios factores, como la es-  
615 tructura de la población, la disponibilidad de un marco muestral y los requerimientos  
616 del estudio. Mientras que un MAS puede implmentarse mediante un procedimiento  
617 sencillo de selección uniforme, un muestreo estratificado requiere dividir la población  
618 en subgrupos y aplicar la selección dentro de cada uno de ellos. Por otro lado, en  
619 el muestreo por conglomerados, las unidades se seleccionan en grupos predefinidos,  
620 en el muestreo sistemático, la selección sigue un patrón regular basado en intervalo  
621 fijo.

622 A continuación, presentaremos los principales algoritmos de selección, deta-  
623 llando sus procedimientos:

### 624 3.5.1. Muestreo aleatorio simple

625 Es la forma más sencilla de muestreo probabilístico. Se obtiene un MAS de  
626 tamaño  $n$  cuando cada posible subconjunto de  $n$  unidades de la población tiene la  
627 misma probabilidad de ser seleccionado como muestra. Al seleccionar una muestra  
628 aleatoria, el investigador está mezclando la población ante de elegir  $n$  unidades. No  
629 es necesario examinar cada miembro de la población por la misma razón que un  
630 técnico de laboratorio no necesita extraer toda la sanbre de una persona para medir  
631 su conteo de glóbulos rojos, pues, la sangre está suficientemente bien mezclada, por

632 lo que cualquier muestra debe ser representativa.

633 **3.5.2. Muestreo estratificado**

634 **3.5.3. Muestreo por conglomerados**

635 **3.5.4. Muestreo sistemático**

## 4. Factores de expansión





637

## CONCLUSIONES

638

1. Conclusión 1.

639

2. Conclusión 2.

640

3. Conclusión 3.



641

## RECOMENDACIONES

642

1. Recomendación 1.

643

2. Recomendación 2.

644

3. Recomendación 3.



## BIBLIOGRAFÍA

- [1] S. Lohr. *Sampling: Design And Analysis*. 2.<sup>a</sup> ed. CRC press/Chapman & Hall/Taylor & Francis Group, Nueva York, 2019.
- [2] R. De la Madrid. The rigged Hilbert space of the free hamiltonian. Consultado en marzo de 2005 en <http://arxiv.org/abs/quant-ph/0210167>.
- [3] J. Escamilla-Castillo. *Topología*. 2.<sup>a</sup> ed. s.e., Guatemala, 1992.
- [4] N. Haaser y J. Sullivan. *Análisis real*. Tr. Ricardo Vinós. Trillas, México, 1978.
- [5] P. Halmos. *Teoría intuitiva de los conjuntos*. 8.<sup>a</sup> ed. Tr. Antonio Martín. Compañía Editorial Continental, S.A., México, 1973.
- [6] F. Kronz. Quantum theory: von Neumann versus Dirac. Consultado en marzo de 2005 en <http://plato.stanford.edu/entries/qt-nvd/>.
- [7] K. Liu, X. Sun, and S.-T. Yau. Goodness of canonical metrics on the moduli space of Riemann surfaces. *Pure Appl. Math. Q.*, **10**(2):223–243, 2014.
- [8] E. Leader and C. Lorcé, The angular momentum controversy: What’s it all about and does it matter?, *Phys. Rept.* **541**, 163 (2014).
- [9] S. Sternberg. Theory of functions of a real variable. Consultado en abril de 2005 en <http://www.math.harvard.edu/~shlomo>.
- [10] Comisión Económica para América Latina y el Caribe (CEPAL), *Diseño y análisis estadístico de las encuestas de hogares de América Latina*, Metodologías de la CEPAL, N° 5 (LC/PUB.2023/14-P), Santiago, 2023.
- [11] A. Gutiérrez, *Estrategias de muestreo: diseño de encuestas y estimación de parámetros*, Ediciones de la U, Bogotá, 2016.