

Entregable 1: Diseño de la Muestra y Variables

A. Diseño de la Muestra

1. Introducción y Comparación con Propuesta 2016

En la primera junta con LICONSA el 31 de agosto de 2017 recibimos un diseño muestral preliminar elaborado en 2016 para esta encuesta. Sin embargo no lo usamos por varios motivos:

1. En el inciso **c. Unidades de muestreo**, se define a los centros de trabajo del PASL como unidades primarias de muestreo. En el inciso **d. Tipo de muestreo**, se menciona que población del PASL se dividirá en dos dominios de estudio: rural y urbano; pero menciona que la estratificación se hará para cada uno de los 33 centros de trabajo. La descripción es confusa ya que se entiende que los dominios son los mismos estratos que se harán para cada uno de los 33 centros de trabajo. En el documento se confunden los conceptos de dominio y estrato.
2. Posteriormente, en el inciso **e. Selección de unidades de muestreo**, se menciona que se realizará un muestreo aleatorio estratificado con afijación proporcional al tamaño del estrato de los hogares, para distribuir la muestra a los estratos o centros de trabajo. En la tabla que aparece bajo el inciso **f. Tamaño de muestra**, se distribuye la muestra de manera proporcional en 33 centros de trabajo LICONSA, es decir que se usan como estratos en la práctica.
3. La dispersión de la muestra es muy alta y por lo mismo es poco eficiente:
 - a. El número de hogares en muestra por lechería que son 5 selecciones con una tasa de respuesta de 80% y 4 entrevistas esperadas efectivas por lechería.
 - b. El número de Entidades Federativas a visitar son las 32 de todo el país.

Esto genera un diseño de una muestra exageradamente dispersa, lo que genera costos excesivos y una muestra poco eficiente.

Probabilística consideró en su propuesta técnica, un diseño con menor dispersión de la muestra, lo que se reflejó en el presupuesto de dicha propuesta. Este diseño es más que suficiente para los objetivos del estudio porque suponiendo un efecto de diseño de 2.0 como menciona la propuesta de 2016 el error muestral es de +/-4% para cada dominio (rural y urbano). Es importante mencionar que dado que se busca representatividad nacional y por dominios urbano y rural, no es necesario cubrir los 32 estados para que la muestra sea eficiente.

El resumen de cambios propuestos en la metodología es el siguiente:

Concepto	Propuesta 2016	Propuesta 2017
Marco Muestral	Listado Hogares Beneficiarios Liconsa agosto 2016	Listado Hogares Beneficiarios julio 2017
Diseño Muestral y Unidades	Probabilístico 3 etapas	Probabilístico 3 etapas
E1: Unidades Primarias de Muestreo	No hay selección, se asignan entrevistas a todos los centros de trabajo con PPT (33 centros)	Selección de 17 Centros de Trabajo con PPT
E2: Unidades Secundarias de Muestreo	Lecherías dentro de cada UPM (selección aleatoria simple)	Lecherías dentro de cada UPM (selección PPT)
E3: Unidades Terciarias de Muestreo	Hogares - Beneficiarios en cada USM (aleatorio simple)	Hogares - Beneficiarios en cada USM (aleatorio simple)
Dominios	Dos dominios: rural y urbano	Dos dominios: rural y urbano. Se realiza un segundo procedimiento de asignación proporcional de las lecherías en cada dominio, utilizando la proporción del número total de hogares considerando solo los 17 centros de trabajo seleccionados en la primera etapa.
Selección UPM: Centros de Trabajo	No hay selección, asignación PPT y se visitan todos (33 centros de trabajo)	Selección 17 centros de trabajo en 5 estratos o regiones PPT del estrato/dominio
Tamaño de Muestra/Selección	2,535 selecciones y se esperan 2,405 entrevistas aplicadas.	3,312 selecciones y se esperan 2,484 entrevistas aplicadas.
Método determinación Tamaño de Muestra	Igual en ambos: A partir de objetivo error muestral, tasa de no respuesta y Efecto de Diseño. Fórmula aplicada con diferencias	
Supuestos DEFF, NR	DEFF es 2.0 y No respuesta 20%	DEFF es 2.0 y No respuesta 25%
Plan de Muestra - Hogares por lechería	Se visitan 507 lecherías con un estimado de 4 entrevistas efectivas por lechería	Se visitan 184 lecherías con un estimado de 13.5 entrevistas efectivas por lechería

2. Objetivo

Generar información estadística sobre la percepción de las personas beneficiarias referente al producto, punto de venta y calidad del servicio del Programa de Abasto Social de Leche en la modalidad de Abasto Comunitario.

3. Descripción inicial

Se realizará muestreo probabilístico en tres etapas con representatividad nacional y para los dominios de zona rural y urbana, dirigido a hogares de personas beneficiarias del programa del PASL en la modalidad de Abasto Comunitario.

4. Población objetivo

La población objetivo la constituyen todos los beneficiarios del programa del PASL en la modalidad de Abasto Comunitario en los 32 estados de la República Mexicana. Al cierre de julio de 2017 el padrón registró un total de 3,333,397 hogares, de los cuales el 69% se localizan en zonas urbanas y el 31% restante en zonas rurales.

5. Marco de muestreo

Como marco muestral se utilizará el registro electrónico del padrón de beneficiarios titulares (PBT) de Liconsa actualizado a julio de 2017, así como el listado de puntos de venta registrados y de lecherías para cada punto de venta.

6. Tamaño de muestra

En vista de que la mayoría de los indicadores se expresan como porcentajes estimados sobre la población beneficiaria, se utiliza el límite máximo para la estimación de proporciones con el criterio de máxima varianza para la determinación del tamaño de muestra. Además, se fija un margen de error absoluto máximo del 4% por dominio, un nivel de confianza de 95%, una tasa de no respuesta aproximada del 25% y un efecto de diseño de 2.0, la fórmula utilizada para el tamaño de muestra es:

$$n = \frac{Z_{\alpha/2}^2 P(1-P) DEFF}{\delta^2 TR}$$

donde

n = Tamaño de muestra

P = Proporción de estimación que produce la máxima varianza (0.5)

$Z_{\alpha/2}$ = Cuantil de una distribución normal asociado a un nivel de confianza del 95%

δ = Margen de error absoluto máximo que se está dispuesto a aceptar (4%)

TNR = Tasa de no respuesta esperada (25%)

$DEFF$ = Efecto de diseño (2.0).

Se obtiene una selección de 1,656 beneficiarios para cada dominio y 3,312 a nivel nacional. Se espera una muestra final efectiva de alrededor de 1,242 entrevistas

concretadas a beneficiarios por dominio por las pérdidas que resultan de la tasa de no respuesta estimada.

7. Plan de muestreo

El diseño de muestreo considera tres etapas de selección:

1. En la primera se utilizarán como unidades primarias (UPM) los 33 centros de trabajo estratificados de manera regional.
2. En la segunda etapa se utilizarán como unidades secundarias de muestreo (USM), los puntos de venta registrados en cada UPM.
3. Finalmente, los hogares registrados en cada USM se considerarán como las unidades terciarias de muestreo (UTM).

La muestra se distribuyó según el plan propuesto en la tabla 1. En esta tabla se establece la distribución de la muestra en las tres etapas, considerando la muestra planeada a partir del cálculo del tamaño y la muestra efectiva esperada al final del levantamiento según la tasa de no respuesta esperada;

Tabla 1. Plan de muestreo por etapa de selección

Etapa	Por lechería	Por dominio	Nacional
Centros de Trabajo (UPM)			17
Lecherías (USM)		92	184
Hogares planeados (UTM)	18	1,656	3,312
Hogares efectivas (UTM)	13.5	1,242	2,484

8. Estratificación

Se utilizará una estratificación regional agrupando los centros de trabajo en cuatro regiones que son:

1. **Región Norte:** Aguascalientes, Baja California, Baja California Sur, Chihuahua, Coahuila, Durango, Nuevo León, San Luis Potosí, Sinaloa, Sonora, Tamaulipas y Zacatecas.
2. **Región Centro:** Colima, Guanajuato, Hidalgo, Jalisco, Michoacán, Morelos, Nayarit, Querétaro y Valle de Toluca.
3. **Región Metropolitana:** Metro Norte y Metro Sur.
4. **Región Sur:** Campeche, Chiapas, Guerrero, Oaxaca, Puebla, Quintana Roo, Tabasco, Tlaxcala, Veracruz y Yucatán.

El objetivo de esta estratificación es agrupar a los beneficiarios de forma que se pueda reducir la varianza dentro de cada estrato y tener estratos que se diferencien entre sí.

Usamos la combinación de dos criterios para generar estos estratos. Primero, la agrupación según la cultura y actitudes de la población en la República Mexicana que están asociadas a la percepción y satisfacción con de los servicios o programas (por ejemplo la diferencia entre el Norte y el Sur).

Segundo, la forma en que LICONSA procesa y distribuye la leche. Por ello generamos el estrato Metropolitano que incluye los centros de trabajo Metro Sur y Metro Norte y representa casi 39% de los hogares beneficiarios. Las diferencias de operación entre estratos son por ejemplo que la zona Norte tiene pocos beneficiarios (menos del 14% del total), sólo se atiende con leche en polvo y toda se atiende desde una planta en Chihuahua. O en otro ejemplo, las zonas metropolitanas y el Valle de Toluca (Estado de México) tienen sus propias plantas de producción cada una y con producto líquido en su mayoría.

En la tabla 2 se muestra el número total de hogares registrados en cada estrato para cada dominio de estudio.

Tabla 2. Distribución del número de hogares por estrato para cada dominio.

Región	Ámbito		Total
	Rural	Urbano	
Zona Norte	145,379	313,474	458,853
Zona Centro	367,871	487,394	855,265
Zona Metropolitana	122,776	1,173,381	1,296,157
Zona Sur	408,955	314,167	723,122
Total	1,044,981	2,288,416	3,333,397

9. Procedimiento de selección

Primera Etapa: Se tomará el listado de los 33 centros de trabajo, de los cuales se seleccionarán los 2 que corresponden al estrato de la Zona Metropolitana y 5 en cada estrato restante, para un total de 17 centros de trabajo en muestra. Se utilizará el procedimiento de selección con Probabilidad Proporcional al Tamaño (PPT), tomando como medida de tamaño el número total de beneficiarios por centro de trabajo.

Segunda etapa: Para la selección de las lecherías por dominio, se realizará una asignación proporcional a los estratos según el número de hogares registrados en el padrón. En la tabla 3 se muestra la distribución proporcional de los hogares registrados en cada estrato para cada dominio de estudio.

Tabla 3. Distribución proporcional de los hogares por estrato para cada dominio.

Región	Ámbito		Total
	Rural	Urbano	
Zona Norte	13.9%	13.7%	13.8%
Zona Centro	35.2%	21.3%	25.7%
Zona Metropolitana	11.7%	51.3%	38.9%
Zona Sur	39.1%	13.7%	21.7%
Total	100.0%	100.0%	100.0%

A partir de ésta distribución, la asignación proporcional de las 92 lecherías en ambos dominios se presenta en la tabla 4.

Tabla 4. Asignación proporcional de la muestra de lecherías por estrato para cada dominio.

Estrato (Región)	Dominio (Ámbito)		Total
	Rural	Urbano	
Zona Norte	13	13	26
Zona Centro	32	19	51
Zona Metropolitana	10	47	57
Zona Sur	37	13	50
Total	92	92	184

Para controlar la dispersión de la muestra y reducir la varianza de los estimadores, se realizará un segundo procedimiento de asignación proporcional de las lecherías en cada dominio, utilizando la proporción del número total de hogares considerando solo los 17 centros de trabajo seleccionados en la primera etapa.

Tercera etapa: Para cada lechería seleccionada, de la base de datos del padrón de beneficiarios titulares (PBT), se seleccionarán los 18 hogares de la muestra planeada con un procedimiento de muestreo aleatorio simple (MAS). Se considera una muestra de remplazo de 4 hogares adicionales, los cuales serán utilizados únicamente cuando la información del padrón no concuerde con la situación actual del hogar visitado (principalmente por falta de actualización o errores de marco).

10. Factores de expansión

El primer paso para el cálculo de los factores de expansión es establecer las probabilidades de selección en cada etapa. Posteriormente, se calcula el factor de expansión inicial, que es el inverso del producto de las probabilidades de selección de cada etapa.

La probabilidad de selección por PPT de las UPM en primera etapa considerando los estratos es la siguiente:

$$p_{hi} = n_h \frac{T_{hi}}{\sum_{i=1}^{N_h} T_{hi}}$$

donde

p_{hi} = Probabilidad de selección de UPM

n_h = Número de UPM en muestra del estrato h -ésimo

N_h = Número de UPM en el estrato h -ésimo

T_h = Número de hogares en la UPM i -ésima del estrato h -ésimo

La probabilidad de selección por PPT de las USM en segunda etapa es la siguiente:

$$p_{hij} = m_{hi} \frac{T_{hij}}{\sum_{j=1}^{M_{hi}} T_{hij}} = m_{hi} \frac{T_{hij}}{T_{hi}}$$

donde

p_{hij} = Probabilidad de selección de USM

m_{hi} = Número de USM en muestra de la UPM i -ésima en el estrato h -ésimo

M_{hi} = Número de USM en la UPM i -ésima del estrato h -ésimo

T_{hij} = Número de hogares en la USM j -ésima en la UPM i -ésima del estrato h -ésimo

La probabilidad de selección por MAS de las USM en tercera etapa es la siguiente:

$$p_{hijk} = \frac{t_{hij}}{T_{hij}}$$

donde

p_{hijk} = Probabilidad de selección de UTM

T_{hij} = Número de UTM en muestra de la USM j -ésima en la UPM i -ésima en el estrato h -ésimo

La expresión del factor de expansión inicial es:

$$w_{hijk} = \frac{1}{p_{hi} \cdot p_{hij} \cdot p_{hijk}} = \frac{\sum_{i=1}^{N_h} T_{hi}}{n_h \cdot T_{hi}} \frac{T_{hi}}{m_{hi} \cdot T_{hij}} \frac{T_{hij}}{t_{hij}} = \frac{\sum_{i=1}^{N_h} T_{hi}}{n_h \cdot m_{hi} \cdot t_{hij}}$$

donde

w_{hijk} = Factor de expansión inicial de la UTM k -ésima en muestra de la USM j -ésima en la UPM i -ésima en el estrato h -ésimo.

El segundo procedimiento consiste en realizar un ajuste por no respuesta de los factores de expansión, es necesario revisar las tasas de no respuesta y evaluar su impacto en las proyecciones poblacionales.

$$w^*_{hijk} = w_{hijk} \frac{t_{hij}}{t^*_{hij}}$$

donde

w^*_{hijk} = Factor de expansión corregido por no respuesta de la UTM k -ésima en muestra de la USM j -ésima en la UPM i -ésima en el estrato h -ésimo.

t^*_{hij} = total de hogares en muestra con respuesta efectiva en la USM j -ésima en la UPM i -ésima en el estrato h -ésimo.

En caso de existir pérdidas de UPM o USM en muestra se realizarán los ajustes correspondientes en cada etapa con procedimientos equivalentes al descrito anteriormente.

En caso de que la no respuesta sea importante, se utilizará un procedimiento de calibración de los factores de expansión, utilizando procedimientos de post-estratificación para variables conocidas de la población y que pueden contribuir a mejorar la representatividad, siempre que se cuente con dicha información.

El aproximado inicial de los factores de expansión a número de beneficiarios que son casi 6.25 millones se muestra a continuación suponiendo probabilidades iguales de selección. Sin embargo, los verdaderos factores de expansión dependen de los tamaños de las unidades primarias (UPM) y secundarias (USM) ya que los procedimientos de selección se hacen probabilidad proporcional al tamaño (PPT) y se estimarán para cada caso en la base una vez que se tenga el número definitivo de entrevistas en cada unidad de muestreo.

Tablas 5 y 6. Estimado Inicial (suponiendo Iguales Probabilidades de Selección) de Factores de Expansión a Total Beneficiarios

Región	Beneficiarios	Muestra planeada	Factor de expansión aproximado*
Zona Norte	835,497	468	1,785
Zona Centro	1,640,176	918	1,787
Zona Metropolitana	2,570,034	1,026	2,505
Zona Sur	1,199,671	900	1,333
Total	6,245,378	3,312	1,886

* El factor de expansión real depende del tamaño de las UPM y USM.

Región	Ámbito				Factor de expansión aproximado por ámbito*	
	Rural		Urbano		Rural	Urbano
	Beneficiarios	Muestra	Beneficiarios	Muestra		
Zona Norte	286,395	234	549,102	234	1,224	2,347
Zona Centro	792,224	576	847,952	342	1,375	2,479
Zona Metropolitana	247,252	180	2,322,782	846	1,374	2,746
Zona Sur	741,361	666	458,310	234	1,113	1,959
Total	2,067,232	1,656	4,178,146	1,656	1,248	2,523

* El factor de expansión real depende del tamaño de las UPM y USM.

11. Bibliografía

- Cochran, W.G., *Sampling Techniques*, 3rd Ed., John Wiley and Sons, New York, 1977.
- Lohr, S. *Sampling: Design and Analysis*, Duxbury Press, Pacific Grove, 1999.
- Raj, D., *Sampling Theory*, and McGraw-Hill, Inc., New York, 1968.

B. Variables

Probabilística recibió en la Junta del 31 de agosto un cuestionario elaborado en 2016 para realizar el levantamiento.

A partir de este cuestionario Probabilística pone a consideración de Liconsa una versión con algunas adaptaciones, dudas y adición de variables. La propuesta de presenta en el entregable 2 y aquí hacemos la lista de variables distinguiendo cuáles venían en la propuesta 2016, cuáles estamos agregando y en cuáles hemos hecha alguna propuesta de cambios pequeños en redacción.

num Pregunta	VARIABLE	ORÍGEN/ ESTATUS	Motivo - Racionalidad
	CLASIFICACIÓN		
	Municipio	Prop2016	
	Centro de trabajo	Prop2016	
	Clave de lechería	Prop2016	
	Clasificación Punto de Venta (tipo)	Prop2016	
	Folio Tarjeta Retiro - Cartón	Prop2016	
	Perfil Beneficiario		
P. Introd	Verificación Titular de tarjeta Liconsa	Prop2016	
P.0A.1	Cantidad de integrantes en su familia	Sugerencia-nuevo	Permiten conocer integrantes, beneficiarios por familia y edades (solicitó SEDESOL)
P.0A.1.1	Número de beneficiarios Liconsa en el hogar	Sugerencia-nuevo	
P.0A.2	Edad de cada uno de los beneficiarios	Sugerencia-nuevo	
P.1	Calificación general a Liconsa (1-10)	Prop2016+ sug. redacción	Se agrega "en general" para asegurar que informante no califique algo específico como la lechería
P.9B	Motivo de compra leche Liconsa	Sugerencia-nuevo	Permitirá segmentar los beneficiarios por motivo de compra
	Calidad en el servicio		
P.2	Tiempo para entrega de tarjeta Liconsa	Prop2016+ sug. redacción	Agregamos ancla tiempo - período entrega documentos (sólo tiempo de Liconsa)
P.3	Realizó solicitud nuevo beneficiario	Prop2016+ sug. redacción	Se acotó al último año
P.4	Tiempo para entrega tarjeta nuevo beneficiario	Prop2016	
P.4B	Recibe información sobre cómo hacer solicitud	Sugerencia-nuevo	Establecer si además de que esté a la vista se da información de incorporación de otra forma
P.5	Lista de requisitos incorporación a la vista en punto venta	Prop2016+ sug. redacción	Facilitar entendimiento de los entrevistados
P.6	Información sobre promotores a la vista en punto de venta	Prop2016+ sug. redacción	
P.7	Conocimiento de los promotores	Prop2016+ sug. redacción	
P.8	Calificación promotores - cortesía y respeto	Prop2016+ sug. redacción	
P.9	Calificación atención general promotores	Prop2016+ sug. redacción	
P.10	Incidencia de quejas Liconsa	Prop2016	

num Pregunta	VARIABLE	ORÍGEN/ ESTATUS	Motivo - Racionalidad
P.11	Incidencia de expresión de quejas	Prop2016	Facilitar entendimiento de los entrevistados
	Medio de Expresión Quejas	Prop2016+ sug. redacción	
	Motivo NO expresión de quejas	Prop2016+ sug. redacción	
P.12	Atención de quejas	Prop2016+ sug. redacción	
P.13	Calificación a la atención de quejas	Prop2016+ sug. redacción	
Producto			
P.14	Calificación general a leche Liconsa	Prop2016+ sug. redacción	Generar variable más dispersión (1 a 10) y centrada en la leche
P.15	Conoce/recibe Información nutricional/fortificación	Prop2016	
P.16	Beneficios nutricionales leche Liconsa (abierta)	Prop2016+ sug. redacción	Facilitar entendimiento de los entrevistados y centrar respuestas de precio
P.17	Precio de compra leche Liconsa	Prop2016+ sug. redacción	
P.18	Percepción - Calificación del precio	Prop2016+ sug. redacción	
P.19	Gasto en leche Liconsa último mes	Prop2016	
P.20	Calificación sabor de leche	Prop2016	
P.21	Disposición a comprar leche Liconsa precio comercial	Prop2016	
P.22	Percepción de ayuda en alimentación/nutrición leche Liconsa	Prop2016+ sug. redacción	Facilitar entendimiento
SECCIÓN LECHE EN POLVO			
P.23	Firma en libro de retiro (compra)	Prop2016+ sug. redacción	Facilitar entendimiento
P.24	Conoce/leyó instrucciones para preparación de leche	Prop2016	
P.25	Claridad de instrucciones	Prop2016+ sug. redacción	Facilitar entendimiento
P.26	Se ha quedado sin dotación de leche (último mes)	Prop2016	
SECCIÓN LECHE LIQUIDA			
P.27	Perforación de tarjeta (compra)	Prop2016	
P.28	Tiempo de entrega	Prop2016+ sug. redacción	Cambiamos tiempo en fila por tiempo de espera porque puede haber más de una fila
SECCIÓN TODOS			
P.28C	Calificación atención concesionarios	Sugerencia-nuevo	Contar con otro elemento para explicar satisfacción beneficiarios
P.28D	Necesidad - Sugerencias de mejora a Liconsa	Prop2016	
P.29	Condicionamiento venta de leche a otras compras	Prop2016+ sug. redacción	Facilitar entendimiento
P.29B	Condicionamiento de venta voto político	Sugerencia-nuevo	Medir incidencia clientelismos político y tener más elementos explicar satisfacción beneficiarios
P.30	Sugerencia para mejorar servicio Liconsa	Prop2016	
Programas Sociales y Verificación Cartilla			
P.I	Beneficiario de otros programas sociales	Prop2016	
P.II	Considera apoyos gobierno como parte del ingreso mensual	Prop2016	
P.III	Recibió cartilla social	Prop2016	
P.IV	Comprensión de información en cartilla social	Prop2016	
P.V	Cartilla social completa (beneficios)	Prop2016	
P.VI	Beneficios recibidos NO incluidos en cartilla	Prop2016	
P.VII	Beneficios NO recibidos incluidos en cartilla	Prop2016	
P.VIII	Conocimiento de monto ahorrado en compra de leche	Prop2016	

