

INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA

**DIRECCIÓN GENERAL DE
ESTADÍSTICAS ECONÓMICAS**

DIRECCIÓN DE MARCOS Y
MUESTREO

Encuesta Mensual de Opinión
Empresarial 2019
(EMOE 2019)

Serie 2013

Comercio

Documento Metodológico
sobre Diseño Muestral



INSTITUTO NACIONAL
DE ESTADÍSTICA Y GEOGRAFÍA

Índice

1. Objetivo de la encuesta	1
2. Población objetivo	1
3. Cobertura sectorial y geográfica	1
4. Dominio de estudio	1
5. Unidad de muestreo	1
6. Marco de muestreo	1
7. Estratificación	1
8. Esquema de muestreo	1
9. Tamaño de la muestra	2
10. Afijación de la muestra	2
11. Selección de la muestra	3
12. Cálculo de los factores de expansión	3
12.1. Ajuste por no respuesta	3
13. Estimadores	4

1. Objetivo de la encuesta

Generar mensualmente indicadores cualitativos de confianza y de tendencia sobre la actividad económica comercial que permita conocer anticipadamente su comportamiento y sirva de apoyo en la toma de decisiones.

2. Población objetivo

La integran las empresas del país pertenecientes al sector comercio según el Sistema de Clasificación Industrial de América del Norte 2013 (SCIAN 2013), con un total de personal ocupado mayor o igual a 50.

3. Cobertura sectorial y geográfica

La cobertura de la encuesta permite obtener estimaciones nacionales para cada sector de actividad comercial integrado por el sector de comercio al por menor y el sector de comercio al por mayor, con base en el SCIAN 2013.

4. Dominio de estudio

El sector comercio al por mayor y comercio al por menor, según el SCIAN 2013 son dominios de estudio.

5. Unidad de muestreo

La unidad de muestreo para la encuesta son las empresas del sector comercio, las cuales pueden estar integradas por un establecimiento comercial o más.

6. Marco de muestreo

El marco se formó por las empresas que contaban con 50 o más personas ocupadas, provenientes del Marco Estadístico Nacional de Unidades Económicas (MENUE), el cual se alimenta del Registro Estadístico de Negocios de México (RENEM) versión octubre 2018 y de los resultados definitivos de los Censos Económicos 2014 (CE 2014).

El marco de muestreo está formado por **6 180** empresas comerciales.

7. Estratificación

Las empresas del marco muestral se agruparon en cinco estratos con base en el total de personal ocupado¹, la distribución resultante se muestra a continuación:

Cuadro 1

Sector comercio	Total	Estratos de Personal Ocupado				
		1	2	3	4	5
		1 001 y más	501 – 1 000	251 - 500	101 - 250	50 - 100
Total	6 180	315	292	650	1 962	2 961
43 Comercio al por mayor	2 710	116	122	277	885	1 310
46 Comercio al por menor	3 470	199	170	373	1 077	1 651

8. Esquema de muestreo

El esquema de muestreo es probabilístico y estratificado.

a) Probabilístico

Las unidades de selección tienen una probabilidad conocida y distinta de cero de ser seleccionadas.

b) Estratificado

Las unidades de muestreo con características similares se agrupan para formar estratos.

¹ Total de Personal Ocupado. Remunerado, no remunerado y suministrado por otra razón social.

9. Tamaño de la muestra

Para el cálculo de tamaño de muestra se utilizó:

- El índice de Confianza del Productor de la Encuesta Mensual de Opinión Empresarial del sector comercio de octubre de 2017 del 49.31% con una varianza de 0.0002.
- El índice de Tendencia Económica de la Encuesta Mensual de Opinión Empresarial del sector comercio de diciembre de 2015 del 58.64% con varianza de 0.0112.

Un error relativo del 4%, un nivel de confianza de 95% y una tasa de no respuesta esperada de 15%.

El tamaño de muestra resultante es de **532** empresas, el cual incluye **315** empresas con más de 1 000 personas ocupadas como certeza.

La expresión para el cálculo del tamaño de la muestra en cada dominio es:

$$n_o = \frac{z^2 \hat{s}^2}{d^2(1 - TNR)} \quad \text{donde: } d = r\hat{l}$$

Aplicando corrección por finitud:

$$n_d = \frac{n_o}{1 + \frac{n_o}{N_d}}$$

Donde:

- n_o = tamaño de muestra sin corrección por finitud.
- z = valor en tablas para una distribución normal estándar.
- \hat{s}^2 = varianza estimada del índice.
- r = nivel de error relativo.
- \hat{l} = índice agregado de tendencia estimado.
- TNR = tasa de no respuesta.
- n_d = tamaño de muestra en el d-ésimo dominio con corrección por finitud.
- N_d = número de empresas en el marco en el d-ésimo dominio.

10. Afijación de la muestra

Para cada dominio con diseño probabilístico, la muestra en cada estrato se asignó por el método de Neyman de acuerdo a la ponderación de personal ocupado por estrato, mediante la siguiente expresión:

$$n_{dh} = n_d \left(\frac{N_{dh} S_{dh}}{\sum_{h=2}^H N_{dh} S_{dh}} \right)$$

Donde:

- n_{dh} = muestra del h-ésimo estrato, del d-ésimo dominio.
- n_d = muestra total del d-ésimo dominio.
- N_{dh} = total de empresas en el h-ésimo estrato, en el d-ésimo dominio.
- S_{dh} = desviación estándar del h-ésimo estrato, del d-ésimo dominio, de la variable del total de personal ocupado.
- H = número de estratos.

La distribución de la muestra se presenta en el siguiente cuadro:

Cuadro 2

Sector comercio	Total	Estratos de Personal Ocupado					Mínima necesaria
		1	2	3	4	5	
		1 001 y más	501 – 1 000	251 - 500	101 - 250	50 - 100	
Total	532	315	51	46	76	44	466
43 Comercio al por mayor	229	116	27	28	34	24	192
46 Comercio al por menor	303	199	24	18	42	20	274

11. Selección de la muestra

Las unidades que integraron el estrato 1 fueron consideradas como certeza en el diseño por tener más de 1 000 personas ocupadas, para el resto de los estratos la selección de la muestra fue aleatoria e independiente para cada uno de ellos.

12. Cálculo de los factores de expansión

El factor de expansión es el inverso de la probabilidad de seleccionar una empresa del h-ésimo estrato del d-ésimo dominio.

La expresión de cálculo es la siguiente:

$$f_{dh} = \frac{N_{dh}}{n_{dh}}$$

Donde:

f_{dh} = factor de expansión del h-ésimo estrato, del d-ésimo dominio.

N_{dh} = total de empresas en el h-ésimo estrato, en el d-ésimo dominio.

n_{dh} = muestra del h-ésimo estrato, del d-ésimo dominio.

12.1. Ajuste por no respuesta

El ajuste del factor de expansión por no respuesta es aplicado cuando no se realizaron todas las entrevistas esperadas y se calculó con base en la expresión:

$$f'_{dh} = f_{dh} \frac{n_{dh}}{n'_{dh}}$$

Donde:

f'_{dh} = factor de expansión con ajuste por no respuesta del h-ésimo estrato, del d-ésimo dominio.

f_{dh} = factor de expansión del h-ésimo estrato, del d-ésimo dominio.

n_{dh} = número de empresas seleccionadas, en el h-ésimo estrato, en el d-ésimo dominio.

n'_{dh} = número de empresas con respuesta, en el h-ésimo estrato, en el d-ésimo dominio.

13. Estimadores

El estimador del total de la característica X para el d-ésimo dominio es:

$$\hat{X} = \sum_{h=1}^H \sum_{i=1}^{n_{dh}} f_{dhi} X_{dhi}$$

Donde:

\hat{X} = estimador del total de la característica X del d-ésimo dominio.

H = número de estratos.

n_{dh} = número de unidades en la muestra en el h-ésimo estrato, en el d-ésimo dominio.

f_{dhi} = factor de expansión del i-ésimo elemento, del h-ésimo estrato, del d-ésimo dominio.

X_{dhi} = característica a estimar reportada en el i-ésimo elemento, en el h-ésimo estrato, en el d-ésimo dominio.

Para la estimación del índice se utiliza el estimador de razón:

$$\hat{R} = \frac{\hat{X}}{\hat{Y}}$$

Donde, \hat{Y} se define en forma análoga a \hat{X} .

Para la evaluación de los errores de muestreo de las estimaciones requeridas se usa el método de Jacknife para la estimación de la varianza, tomando como referencia la siguiente fórmula para estimar la precisión de $\hat{\theta}$ para cada dominio:

$$\hat{V}_{JK}(\hat{\theta}) = \sum_{h=1}^H \frac{n_h - 1}{n_h} \sum_{j=1}^{n_h} [\hat{\theta}_{(hj)} - \hat{\theta}]^2$$

Donde:

$\hat{V}_{JK}(\hat{\theta})$ = varianza Jacknife del estimador $\hat{\theta}$.

H = número de estratos.

n_h = número de elementos en el h-ésimo estrato.

$\hat{\theta}_{(hj)}$ = estimador en la misma forma que $\hat{\theta}$ pero sin utilizar la j-ésima observación, del h-ésimo estrato.

$\hat{\theta}$ = parámetro a estimar.

Las estimaciones del error estándar (EE) y el coeficiente de variación (CV), se calculan mediante las siguientes expresiones:

$$EE = \sqrt{\hat{V}_{JK}(\hat{\theta})} \quad CV = \frac{\sqrt{\hat{V}_{JK}(\hat{\theta})}}{\hat{\theta}}$$