

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 Construcción

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	3

1. Objetivo de la encuesta

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

2. Población objetivo

La población objetivo la integran las empresas constructoras del país según el Sistema de Clasificación de América del Norte 2013 (SCIAN 2013), con un total de personal ocupado mayor a 100.

3. Cobertura sectorial y geográfica

La cobertura de la encuesta permite obtener estimaciones nacionales para el sector construcción, con base en el SCIAN 2013.

4. Dominio de estudio

El dominio de estudio es considerado a nivel sector, para este caso es el sector construcción según el SCIAN 2013.

5. Unidad de muestreo

La unidad de muestreo para la encuesta son las empresas del sector construcción.

6. Marco de muestreo

El marco se formó por las empresas que contaban con más de 100 personas ocupadas, provenientes del Marco Estadístico Nacional de Unidades Económicas (MENUE), el cual se alimenta del Registro Estadístico de Negocios (RENEM) versión octubre 2018 y de los resultados definitivos de los Censos Económicos 2014 (CE 2014).

El marco de muestreo está formado por **1 323** empresas constructoras.

7. Estratificación

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

Cuadro 1

Estratos	Rango del Total de Personal Ocupado	Número de empresas
Total		1 323
1	1 001 y más	63
2	501 a 1 000	84
3	251 a 500	253
4	101 a 250	923

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 construcción de diciembre de 2015 del 48.94% con una varianza de 0.0002.
- El índice de Tendencia Económica de la Encuesta Mensual de Opinión Empresarial del sector construcción de diciembre de 2015 del 54.92% con varianza de 0.0117.

Un error relativo del 3%, un nivel de confianza de 95% y una tasa de no respuesta esperada de 15%.

El tamaño de muestra resultante es de **239** empresas, el cual incluye **63** 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:

$$n = \frac{n_o}{1 + \frac{n_o}{N}}$$

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.
- TNR = tasa de no respuesta.
- n = tamaño de muestra con corrección por finitud.
- N = número de empresas en el marco.

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:

$$n_h = n \left(\frac{N_h S_h}{\sum_{h=2}^H N_h S_h} \right)$$

Donde:

- n_h = muestra del h-ésimo estrato.
- n = muestra total.
- N_h = total de empresas en el h-ésimo estrato.
- S_h = desviación estándar del h-ésimo estrato, de la variable del total de personal ocupado.
- H = número de estratos.

La distribución de la muestra por estrato se presenta enseguida:

Cuadro 2

Estratos	Rango del Total de Personal Ocupado	Número de empresas	Mínima necesaria
Total		239	206
1	1 001 y más	63	63
2	501 a 1 000	35	25
3	251 a 500	48	39
4	101 a 250	93	79

11. Selección de la muestra

Las unidades que integraron el estrato 1 fueron consideradas como certeza 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.

La expresión de cálculo es la siguiente:

$$f_h = \frac{N_h}{n_h}$$

Donde:

f_h = factor de expansión del h-ésimo estrato.

N_h = total de empresas en el h-ésimo estrato.

n_h = muestra del h-ésimo estrato.

12.1. Ajuste por no respuesta

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

$$f'_h = f_h \frac{n_h}{n'_h}$$

Donde:

f'_h = factor de expansión con ajuste por no respuesta del h-ésimo estrato.

f_h = factor de expansión del h-ésimo estrato.

n_h = número de empresas seleccionadas, en el h-ésimo estrato.

n'_h = número de empresas con respuesta, en el h-ésimo estrato.

13. Estimadores

El estimador del total de la característica X es:

$$\hat{X} = \sum_{h=1}^H \sum_{i=1}^{n_h} f_{hi} X_{hi}$$

Donde:

- \hat{X} = estimador del total de la característica X.
- H = número de estratos.
- n_h = número de unidades en la muestra, en el h-ésimo estrato.
- f_{hi} = factor de expansión del i-ésimo elemento, del h-ésimo estrato.
- X_{hi} = característica a estimar reportada en el i-ésimo elemento, en el h-ésimo estrato.

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 Jackknife para la estimación de la varianza, tomando como referencia la siguiente fórmula para estimar la precisión de $\hat{\theta}$ de 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 Jackknife 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:

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