

MANUAL DE ANÁLISIS ESTADÍSTICO DE DATOS

Autor:

CENTRO DE DERECHOS ECONÓMICOS SOCIALES Y CULTURALES DEL ECUADOR –
CDES

Técnicos responsables:

Irma Jaya
Pablo Iturralde

QUITO, FEBRERO DE 2022

INTRODUCCIÓN

El presente Manual de análisis estadístico de datos es una guía paso a paso de la construcción de indicadores laborales y, la incidencia de pobreza y extrema pobreza. En él se proporcionan instrucciones detalladas y sencillas para reunir y procesar datos, construir estimadores estadísticos e indicadores a fin de que la nueva información obtenida permita examinar la situación del país. El Manual contribuirá a crear capacidad estadística y a impulsar la producción estadística.

CARACTERIZACIÓN DE LA POBLACIÓN ECONÓMICAMENTE ACTIVA (PEA).

La revisión de la situación laboral requiere de un análisis de las condiciones demográficas, económicas y sociales (Elizaga & Mellon, 1971); determinantes para que una persona forme parte de la Población Económicamente Activa (PEA). En relación con eso, la importancia y pertinencia del estudio de las características mencionadas, que establecerán las razones para encontrarse dentro de los estados de ocupación laboral: Empleo Adecuado/Pleno, Subempleo (Subempleo por insuficiencia de tiempo de trabajo y Subempleo por insuficiencia de ingresos), Otro empleo no pleno, y Desempleo (Desempleo abierto y Desempleo oculto).

Fuentes de datos

Los datos de insumo para el proceso de producción estadística los conseguimos de la base de datos de la Encuesta Nacional Empleo, Desempleo y Subempleo (ENEMDU), publicada por el Instituto Ecuatoriano de Estadísticas y Censos (INEC), para el periodo 2016-2021. Los datos se analizan a través del paquete estadístico SPSS.

Análisis Descriptivo

Iniciamos con la recodificación de las variables Etnia, Tipo de Trabajo y Rama de actividad económica, obteniendo nuevas categorías, para presentar estas variables como lo hace en sus publicaciones el INEC:

Etnia:

Afrodescendiente: Afroecuatoriano, Mulato y Negro.

Tipo de trabajo:

Empleo Asalariado: Empleados/Obrero de Gobierno/Estado, Empleado/Obrero privado, Empleado/Obrero Tercerizado, Jornalero o Peón, y Empleado(o) Doméstico(a).

No remunerados: Trabajador del Hogar no Remunerado, Trabajador No Remunerado en otro Hogar, y Ayudante no Remunerado de asalariados.

Rama de actividad, CIU 4.0:

Suministro de electricidad y agua: Suministros de electricidad, gas, aire acondicionado; y Distribución de agua, alcantarillado.

Actividades profesionales, técnicas y administrativas: Actividades profesionales, científicas y técnicas; y Actividades y servicios administrativos y de apoyo.

Enseñanza y Servicios sociales y de salud: Enseñanza; y Actividades, servicios sociales y de salud.

Otros Servicios: Actividades inmobiliarias; Artes, entretenimiento y recreación; Otras actividades de servicios; Actividades de organizaciones extraterritoriales; y No especificado.

Para la caracterización de la Población Económicamente Activa (PEA) generamos tablas personalizadas de la PEA por Sexo, Etnia, Tipo de Trabajo y Rama de actividad económica, consiguiendo los recuentos; y aplicamos el ponderador del factor de expansión a la muestra para obtener estimaciones poblacionales. Después, expresamos las categorías de la Población Económicamente Activa —Empleo Adecuado/Pleno, Subempleo, Otro empleo no pleno, y Desempleo— como porcentaje respecto a la PEA.

DISTRIBUCIÓN - CONCENTRACIÓN DEL INGRESO LABORAL E INGRESO PER CÁPITA POR DECILES

Existen distintas maneras de estratificar a una población según sus modos de consumo, de igual modo por sus ingresos y condiciones de trabajo. En tal sentido, ninguna es más válida que otra, cada una permite llegar a distintas conclusiones.

Los deciles de ingreso laboral e ingreso per cápita permitirán ordenar a un conjunto de personas —Población total, Población Ocupada o con empleo, y Población Económicamente Activa (PEA)—, desde los que menos renta hasta los que más renta percibe en diez partes iguales. Así determinar la distribución, concentración y, participación por Sexo, Etnia y Tipo de Trabajo para cada decil de ingreso; además de calcular estimaciones de la variabilidad —promedio, probabilidad, error estándar, intervalo de confianza, relación con el decil 10, brecha, y relación Umbrales Mínimos de Satisfacción - Deciles— para cada decil de ingreso.

Fuentes de datos

Los datos de insumo para el proceso de producción estadística los conseguimos de la base de datos de la ENEMDU, publicada por el INEC; además, fue necesaria la información estadística de la Canasta Familiar Básica y Canasta Familiar Vital, publicada por el INEC, y de la Información Estadística Mensual del Banco Central del Ecuador, para el periodo 2016-2021. Los datos se analizan a través de los paquetes estadísticos SPSS y Stata.

Tratamiento/depuración de la base de datos

Antes de realizar alguna operación estadística, se debe estudiar las variables en busca de indicios de algunos de los siguientes problemas de valores perdidos, atípicos, incompletos, duplicados o conflictos en instancias, entre otros. (Espinosa & al., 2010)

- **Ingreso laboral:**

A los valores aplicables no válidos (-1 = gasta más de lo gana), a la ausencia de respuesta (999999 = No informa), y los datos faltantes les asignamos el valor de 0, pues la mayoría corresponde a personas que no perciben ingresos por su actividad económica. (Castillo & Rosero, 2015)

En relación con los valores atípicos, valores extremos demasiados grandes o pequeños, estos no serán descartados para el análisis por deciles, su eliminación podría conducir a la pérdida de información importante.

- **Ingreso per cápita**

La técnica de imputación de los datos faltantes es la asignación del valor 0, ya que la mayoría corresponde a personas que no perciben ingresos por su actividad económica (Castillo & Rosero, 2015). Los valores atípicos/extremos se conservaban para el análisis por deciles, manteniendo así la mayor cantidad de información posible.

Correcciones muestrales

La ENEMDU tiene un diseño muestral complejo —muestreo probabilístico estratificado bietápico; la Unidad Primaria de Muestreo (UPM) es el conglomerado y la Unidad Secundaria de Muestreo (USM) son las viviendas ocupadas— (INEC, 2018b); tomar en cuenta estas características permite obtener las mejores estimaciones. La información del diseño muestral complejo aplicado en la operación estadística es necesaria para realizar un

correcto cálculo de las estimaciones de la variabilidad; además, permite obtener los mejores estimadores sin exagerar la significancia estadística. (INEC, 2018a)

Para la estimación de los parámetros de interés y sus correspondientes estimaciones de la variabilidad, utilizaremos el programa estadístico Stata, ya que este nos permite aplicar el método de último conglomerado con la linealización de la serie Taylor planteada por el INEC, para realizar las correcciones muestrales.

Análisis Descriptivo

Antes de iniciar el análisis descriptivo por deciles de ingreso, construimos los deciles sobre un determinado conjunto de personas —Población total, Población Ocupada/con empleo, o Población Económicamente Activa (PEA)—, con la ayuda del programa estadístico SPSS. Y continuación obtenemos la distribución, concentración, estimaciones de la variabilidad y participación para cada decil de ingreso; indicando el diseño complejo de la encuesta con la ayuda del programa estadístico Stata.

Estimaciones de la variabilidad para los deciles.

Obtenemos el promedio/media, probabilidad, error estándar, e intervalo de confianza para cada decil de ingreso, con la ayuda de los comandos en Stata.

Establecemos la relación de cada decil de ingreso con el último decil respecto al promedio/media de ingreso, dividiendo el valor del ingreso promedio de cada decil para el valor del ingreso promedio del último decil. Determinamos la brecha del ingreso promedio en términos monetarios y porcentuales para cada decil. Finalmente, expresamos la relación de los Umbrales Mínimos de Satisfacción y deciles, señalando los deciles en términos porcentuales que alcanzan un umbral.

Distribución - Concentración del Ingreso laboral e Ingreso per cápita por deciles.

Para determinar la distribución por decil de ingreso, generamos tablas cruzadas de la población de interés —Población total, Población Ocupada/con empleo, o Población Económicamente Activa (PEA)— por deciles de ingreso; consiguiendo el recuento y porcentaje por columna. Después, para determinar la concentración del ingreso para cada decil, multiplicamos el ingreso promedio de cada decil por los recuentos del respectivo decil de ingreso.

Características de los deciles por Sexo, Etnia y Tipo de Trabajo.

Para determinar la participación por Sexo, Etnia, y Tipo de Trabajo para cada decil, generamos tablas cruzadas de las variables mencionadas por deciles de ingreso, consiguiendo el porcentaje por columna.

POBREZA Y EXTREMA POBREZA POR INGRESOS

La pobreza y extrema pobreza tienen un carácter multidimensional, no obstante hay ciertas dimensiones básicas de estudio como la dimensión sectorial, territorial, familiar y factores adscritos (Arriagada, 2005), que deberán considerarse para una adecuada observación de la problemática.

Fuentes de datos

Los datos de insumo los conseguimos de la base de datos de la ENEMDU, publicada por el INEC, para el periodo 2016-2021; los datos se analizan a través del paquete estadístico SPSS. Las variables de *pobreza* y *extrema pobreza* solo se calculan y presentan en la base de datos de los meses de junio y diciembre.

Análisis Descriptivo

Iniciamos con la recodificación de la variable Seguro para obtener la nueva categoría *Otro Seguro*, como lo hace el INEC:

Otro Seguro: IESS, seguro general voluntario; IESS, seguro campesino; ISSFA, ISSPOL, Seguro privado con hospitalización, Seguro privado sin hospitalización, AUS, Seguros Municipales y Seguro Ministerio Salud Pública.

Incidencia/tasa

Para determinar la incidencia o tasa de pobreza y extrema pobreza a nivel nacional, obtenemos una tabla de la distribución de frecuencias, tomando el porcentaje válido de pobres e indigentes —pobres extremos—. Además, establecemos la incidencia o tasa de pobreza y extrema pobreza por Sexo, Etnia, Tipo de Trabajo, Rama de actividad económica y Seguro, generando tablas personalizadas de la pobreza y extrema pobreza por las características mencionadas; consiguiendo el porcentaje válido por columna, y aplicamos el ponderador del factor de expansión a la muestra para obtener estimaciones poblacionales.

Referencias

- Arriagada, I. (2005). Dimensiones de la pobreza y políticas desde una perspectiva de género. *División de Desarrollo Social, CEPAL*. Obtenido de https://repositorio.cepal.org/bitstream/handle/11362/11002/085101113_es.pdf?sequence=1&isAllowed=y
- Banco Central del Ecuador. (2022). *Información Estadística Mensual*. Obtenido de <https://contenido.bce.fin.ec/home1/estadisticas/bolmensual/IEMensual.jsp>
- Castillo, R., & Rosero, J. (2015). Empleo y condición de actividad en Ecuador. *Revista de Estadística y Metodologías, Instituto Nacional de Estadística y Censos (CGTIMA-INEC), Volumen 1*. Obtenido de https://www.ecuadorencifras.gob.ec/documentos/web-inec/Estudios%20e%20Investigaciones/Trabajo_empleo/3.%20REM-Empleo_conduct.pdf
- Elizaga, J. C., & Mellon, R. (1971). *Aspectos demográficos de la mano de obra en América Latina*. Obtenido de <https://repositorio.cepal.org/handle/11362/7428>
- Encuesta Nacional de Empleo, Desempleo y Subempleo (ENEMDU). (2021). Boletín Técnico N° 02-2021. Obtenido de <https://www.ecuadorencifras.gob.ec/documentos/web-inec/POBREZA/2020/Diciembre-2020/Boletin%20tecnico%20pobreza%20diciembre%202020.pdf>
- Espinosa, R., & al., e. (2010). Hacia la consideración de aspectos de calidad de datos en procesos de minería: el caso de las técnicas de clasificación. *Actas de los Talleres de las Jornadas de Ingeniería del Software y Bases de Datos, Vol. 4*. Obtenido de <http://hdl.handle.net/10045/25165>
- INEC. (2018a). *Cálculo del error de muestreo y declaración de muestras complejas en la Encuesta Nacional de Empleo, Desempleo y Subempleo (ENEMDU)*. Instituto Nacional de Estadística y Censos. Quito-Ecuador. Obtenido de https://www.ecuadorencifras.gob.ec/documentos/web-inec/EMPLEO/2020/Diciembre-2020/202012_Calculo%20de%20errores%20estandar%20y%20declaracion%20de%20muestras%20complejas_ENEMDU.pdf
- INEC. (2018b). *Documento Metodológico de la Encuesta de Empleo, Desempleo y Subempleo (ENEMDU)*. Quito-Ecuador. Obtenido de https://www.ecuadorencifras.gob.ec/documentos/web-inec/EMPLEO/2018/Septiembre-2018/ENEMDU_Metodologia%20Encuesta%20Nacional%20de%20Empleo%20Desempleo%20y%20Subempleo.pdf

INEC. (2019). *Encuesta Nacional Empleo, Desempleo y Subempleo (ENEMDU)*. Obtenido de <https://www.ecuadorencifras.gob.ec/enemdu-diciembre-2019/>

INEC. (2021a). *Canasta*. Obtenido de <https://www.ecuadorencifras.gob.ec/canasta/>

INEC. (2021b). *Pobreza por ingresos*. Obtenido de <https://www.ecuadorencifras.gob.ec/pobreza-diciembre-2021/>