



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO
FACULTAD DE ECONOMÍA

MAESTRÍA EN ESTUDIOS SUSTENTABLES REGIONALES
Y METROPOLITANOS

CUADERNO DE EJERCICIOS DE LA
UNIDAD DE APRENDIZAJE: ESTADÍSTICA APLICADA I

Elaborado por: Ricardo Rodríguez Marcial

Septiembre de 2019

Tabla de contenido

INTRODUCCIÓN	6
ESTADÍSTICA DESCRIPTIVA	8
CONCEPTOS ESTADÍSTICOS BÁSICOS	8
TABULACIÓN Y REPRESENTACIÓN GRÁFICA DE LA INFORMACIÓN	11
MEDIDAS DE POSICIÓN Y DISPERSIÓN	16
ANÁLISIS DE CONCENTRACIÓN.....	20
ESTADÍSTICA INFERENCIAL	21
PRUEBA DE HIPÓTESIS.....	21
ANÁLISIS DE REGRESIÓN	23
BIBLIOGRAFÍA	30

Unidad de aprendizaje: **ESTADISTICA APLICADA I**

CLAVE	PERIODO LECTIVO	HT	HP=	TH=	CRÉDITOS
	1°	2	2	4	6

Elaboró: **Dra. María del Carmen Salgado Vega**

Fecha: Agosto 2010

Propósito:

Contribuir a la formación genérica del estudiante y fortalecer sus competencias disciplinarias en el campo de la estadística. En consecuencia, el alumno dominará los principios de la estadística descriptiva e inferencial y su aplicación a los sistemas de Información geográfica, fenómenos regionales, metropolitanos y sustentables

Contenido Temático y Sintético:

Modulo I

Nivel conceptual y teórico de la estadística descriptiva e inferencial

Modulo II

Aplicación a la economía del desarrollo regional y metropolitanos sustentables de la estadística básica.

Modulo III

Métodos y técnicas de la estadística básica

Modulo IV

Ejercicios aplicados a los sistemas de información geográfica

EVALUACIÓN

Lecturas controladas (10%) y exposición de trabajos (15%).....25 %

Examen parcial.....30%

Trabajos escritos.....45%

BIBLIOGRAFÍA

- Box., G. William H. y Hunter.S. (2008) *Estadística para investigadores*. México. Editorial Reventé S. A.
- García, A. y Simón de Blas, C. (2007). *Manual de Estadística*. Madrid, España. Editorial DYKINSON.
- Levin., R y Rubin. D. (2004). *Estadística para administración y economía*. México. Pearson Prentice Hall.
- Rogerson, P. (2001). *Statistical Methods for Geography*. SAGE Publications. Printed in Great Britain.
- Mendenhall,W. y Reinmuth, J. (2005). *Estadística para administración y economía*. México. Grupo Editorial Iberoamérica.

ESTRUCTURA CURRICULAR

Periodo lectivo			
Primero	Segundo	Tercero	Cuarto
Ciudades complejas y aglomeradas	Optativa básica A	Optativa intermedia A	Trabajo terminal de grado IV
Economía regional	Optativa básica B	Optativa intermedia B	Temas selectos
Responsabilidad social y desarrollo sustentable	Estadística aplicada II	Trabajo terminal de grado III	
Estadística aplicada I	Trabajo terminal de grado II		
Trabajo terminal de grado I			

INTRODUCCIÓN

El Programa de la Unidad de Aprendizaje denominado Estadística Aplicada I es parte fundamental del Programa Interinstitucional de la Maestría en Estudios Sustentables Regionales y Metropolitanos, que se imparte en la Facultad de Economía de la Universidad Autónoma del Estado de México.

Este cuaderno de ejercicios pretende contribuir a que los alumnos, que provienen de distintas formaciones profesionales, se familiaricen con las técnicas estadísticas y les permita mejorar la comprensión de los fenómenos propios del programa de Maestría. De ahí que los ejercicios ocupen información de los ambientes donde se desarrollarán en el ejercicio de su profesión. Se parte desde mostrar la utilidad de los conceptos básicos en la comprensión de los fenómenos, hasta el uso de métodos de estimación donde forma parte sustancial la explicación de los fenómenos, elemento central del método científico.

La importancia del análisis estadístico radica en que las personas y las instituciones, en general la sociedad entera, se encuentra inmersa en un entorno incierto, donde se requiere un conocimiento completo de los fenómenos a estudiar, requiere saber sobre la recolección y análisis de la información estadística existente, que redunde en la toma de decisiones óptima por parte de los agentes decisores.

El método científico se basa en dos tipos de razonamientos: el deductivo y el inductivo. El primero de ellos hace referencia al proceso de partir de lo general a lo particular, mientras que el segundo lo hace en sentido inverso. La investigación Estadística ocupa ambos procesos.

La Unidad de Aprendizaje Estadística Aplicada I, otorgará al alumno la posibilidad de conocer los elementos teóricos y metodológicos necesarios para que distinga entre estadística descriptiva e inferencial y, la utilidad que cada una de las ramas aportan al trabajo profesional. La Estadística DESCRIPTIVA trata de la descripción numérica de conjuntos, incluye técnicas que se relacionan con el resumen y la descripción de datos numéricos. De manera concreta, la estadística descriptiva está formada por los métodos gráficos que se utilizan para resumir y procesar los datos

y transformarlos en información. La Estadística INFERENCIAL se ocupa de interpretar los resultados obtenidos con las técnicas descriptivas, para tomar decisiones con base en los resultados, es decir, se toman decisiones sobre una población estadística basada en una muestra. Es la base para hacer predicciones, previsiones y estimaciones que se utilizan para transformar la información en conocimiento.

ESTADÍSTICA DESCRIPTIVA

CONCEPTOS ESTADÍSTICOS BÁSICOS

1. Especificar la población de donde se extrae la muestra que debería considerarse para realizar un estudio estadístico sobre:
 - a) La valoración económica del servicio ambiental, por el método de valoración contingente, que proporciona el bosque en el Municipio de Valle de Bravo, Estado de México.
 - b) La valoración económica, por el método de costo de viaje, que proporciona el bosque en el Municipio de Valle de Bravo, Estado de México.
 - c) Nivel de satisfacción de los usuarios que proporciona el servicio de transporte universitario POTROBUS.

Respuesta:

- a) Debido a que el método de valoración contingente de servicios ambientales, consiste en que los habitantes de la región, donde se localiza el bosque, valoren económicamente su servicio; la población sería todos los habitantes del Municipio de Valle de Bravo; considerando una muestra de ellos.
 - b) Debido a que el método de valoración por costo de viaje de servicios ambientales, consiste en que las personas que visitan el bosque determine lo que están dispuestos a pagar por disfrutar del servicio; la población de donde se extraería la muestra sería el número de visitantes que recibe la región en un tiempo determinado.
 - c) La población estaría constituida por todos los universitarios usuarios del servicio de transporte denominado POTROBUS.
-
2. Indicar para las características siguientes, qué tipo de variables son:
 - a) Cantidad de dinero que se está dispuesto a pagar por los servicios ambientales que proporciona el bosque.
 - b) Nivel de escolaridad de los habitantes de una región

- c) Problemática de las zonas boscosas
- d) Ocupación de los habitantes de una región

Respuesta:

- a) Cantidad de dinero: cuantitativa
- b) Escolaridad: cualitativa
- c) Problema: cualitativa
- d) Ocupación: cualitativa

3. Señalar la población, variables y tipo de variables, correspondientes a las siguientes situaciones:

- a) Estudio sobre el nivel de satisfacción de los usuarios del servicio de transporte universitario POTROBUS.
- b) Estudio para determinar el valor económico de los servicios ambientales proporcionados por el bosque en una región determinada.
- c) Reclamaciones por los servicios prestados por bancos recibidas en la Comisión Nacional para la Protección y Defensa de los Usuarios de Servicios Financieros (CONDUSEF).

Respuesta:

- a) Población: Usuarios del servicio de transporte universitario.
Variable: Nivel de satisfacción.
Tipo de variable: Cualitativa
- b) Población: Población de la región de estudio.
Variable: Valor del servicio ambiental
Tipo de variable: Cualitativa
- c) Población: Consumidores de los servicios bancarios
Variable: Número de reclamaciones
Tipo de reclamación
Tipo de variable: Cuantitativa
Cualitativa

4. Un grupo de investigadores que quiere determinar el valor económico de los servicios ambientales que ofrece el bosque en el Municipio de Valle de Bravo en el Estado de México, ¿Qué información deberían recoger? Se pide identificar las características más relevantes para la realización de la tarea y diferenciarlas según sean cuantitativas o cualitativas.

Respuesta:

Entre las características más relevantes, destacan:

- Localidades que integran el Municipio: Cualitativa
- Número de personas: cuantitativa
- Género: masculino femenino. Cualitativa
- Edad. Cuantitativa
- Escolaridad. Cualitativa
- Cualidades que aprecian del bosque. Cualitativa
- Ingreso mensual. Cuantitativa
- Cantidad dispuesta a pagar. Cuantitativa

Esta información se recogería a través de una encuesta, se estimaría una muestra y se dividiría en cada una de las localidades seleccionadas para el estudio.

5. Como parte de un estudio sobre el tamaño de establecimientos según el número de empleados en el Estado de México, se recogió información de la zona metropolitana del Municipio de Toluca; los resultados muestrales para esta zona se representan en la Tabla 1, en número de establecimientos:

Tabla 1.

	Metepec	Toluca	Zinacantepec	Lerma
De 11 a 30 empleados	18	104	5	67
De 31 a 50 empleados	6	29	2	30

Respuesta:

Población: Número de establecimientos con 11 a 30 empleados y de 31 a 50 empleados

Muestra: Se elige una muestra representativa

Variable: Municipios

Número de empleados

Tipo de variables: Cualitativa

Cuantitativa

Escala de medida: Nominal

TABULACIÓN Y REPRESENTACIÓN GRÁFICA DE LA INFORMACIÓN

6. Se hizo un estudio por parte de la Facultad de Economía para determinar la cantidad de dinero que estarían dispuestos a pagar los encuestados por los servicios ambientales que presta el bosque. Se eligió una muestra de 265 personas cuyos salarios son los siguientes:

193 personas tienen un salario menor a 3,000 pesos

0 personas tienen un salario entre 3,000 y 6,000 pesos

41 personas tienen un salario entre 6,000 y 9,000 pesos

31 personas tienen un salario entre 9,000 y 12,000 pesos

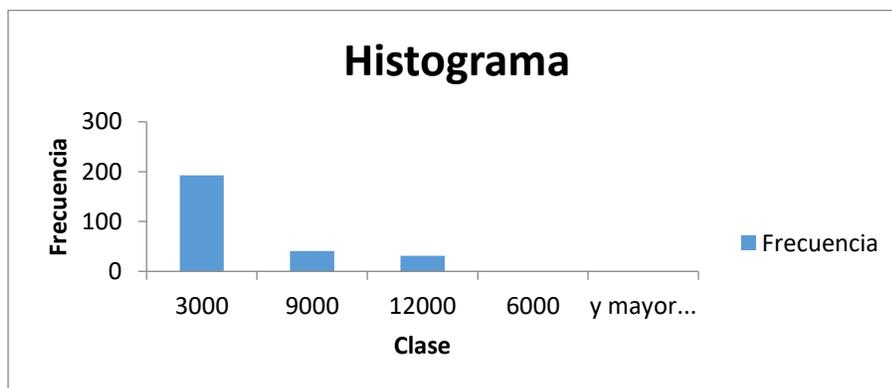
A partir de esta información

- Especificar la población, la variable observada y el tipo de variable.
- Construir la representación gráfica adecuada
- Obtener las frecuencias absolutas acumuladas y representarlas gráficamente
- Determinar la proporción de población muestral que perciben ingresos superiores a los 6,000 pesos.

Respuesta:

- a. La población serán todos los ciudadanos trabajadores del Municipio de Valle de Bravo, Estado de México. La variable observada es el salario, y el tipo de variable cuantitativa continua.
- b. Para construir la gráfica, previamente se tabulan los datos. Teniendo en cuenta el tipo de variable, la representación gráfica adecuada es el histograma.

Salarios	n_i	N_i	f_i	F_i
Menor a 3,000	193	193	0.728	0.728
3,000 – 6,000	0	193	0	0.728
6,000 – 9,000	41	234	0.155	0.883
9,000 – 12,000	31	265	0.117	1.000



- c. La representación gráfica para las frecuencias absolutas acumuladas se obtiene a partir del histograma y se expresa:



- d. La proporción de trabajadores con ingresos superiores a los 6,000 'esos será:
 $0.155 + 0.117 = 0.372$, es decir. El 37.2%.

7. La Facultad de Economía ha recopilado información relativa a las personas por nivel de escolaridad, según sexo y edad. Los datos para el año 2018 se expresan, en número de personas:

	Hombres				Mujeres			
	20-30	31-40	41-50	51 a más	20-30	31-40	41-50	51 a más
Sin estudio	1	3	10	19	0	5	4	5
Primaria	3	8	22	9	1	3	5	2
Secundaria	10	17	13	2	10	14	5	3
Preparatoria	0	0	0	0	10	8	0	0
Universidad	5	5	9	3	4	2	0	0

- Especificar la población, las variables consideradas y el tipo de variable.
- Calcular el porcentaje de hombres para cada uno de los niveles de escolaridad.
- Obtener los porcentajes de mujeres que corresponden a cada grupo de edad.
- Representar gráficamente la distribución de hombres por edad y las mujeres por escolaridad.

Respuesta:

- Población: Trabajadores del Municipio de Valle de Bravo.

Variables: Sexo: Cualitativa.

Edad: Cuantitativa.

Escolaridad: Cualitativa.

- Porcentaje de hombres por Escolaridad

Escolaridad	Hombres (n _i)	% Hombres
Sin estudios	33	23.7
Primaria	42	30.2
Secundaria	42	30.2
Preparatoria	0	0.0
Universidad	22	15.9
Total	139	

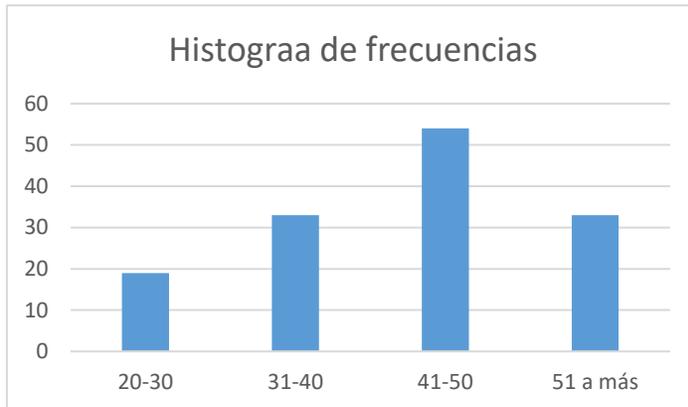
- Porcentaje de mujeres por grupo de edad

Edad	Mujeres (n _i)	% Mujeres
20 - 30	25	30.9
31 - 40	32	39.5
41 - 50	14	17.3
51 a más	10	12.3
Total	81	

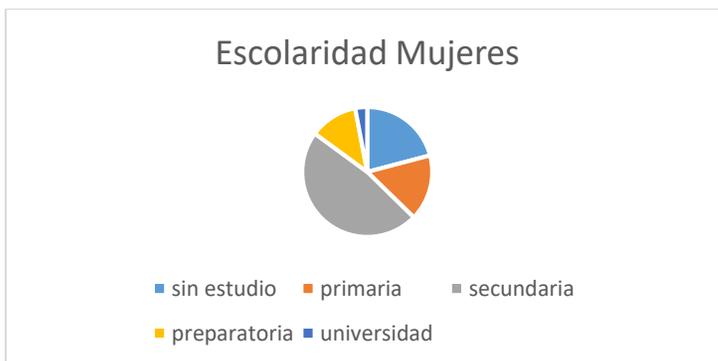
- Representación gráfica

Edad	Hombres
21 - 30	19
31 - 40	33
41 - 50	54

51 a más	33
----------	----



Escolaridad	Mujeres
Sin estudios	14
Primaria	11
Secundaria	32
Preparatoria	8
Universidad	2



8. Se recogió información sobre los ingresos de personas del municipio de Valle de Bravo, clasificados según su estado civil:

Estado civil	Salario		
	2,000 – 4,000	5,000 – 8,000	9,000 -12,000
Soltero	24	0	3
Casado	149	30	14
Unión libre	21	9	0

A partir de esta información, y una vez realizado el diseño del estudio estadístico, se pide:

a. Porcentaje de trabajadores casados que cobran más de 3,000 pesos.

- b. Realizar el estudio comparativo del salario de los trabajadores con estado civil soltero y en unión libre.

Respuesta:

Diseño del estudio

Población: Trabajadores del Municipio de Valle de Bravo

Muestra: Trabajadores que opinaron sobre el valor del servicio ambiental del bosque.

Variables: Estado civil: Cualitativa

Salario: Cuantitativa

- a. Trabajadores casados

Salario	n_i	N_i
2,000 – 4,000	149	149
5,000 – 8,000	30	179
9,000 – 12,000	14	193

Se trata de calcular el valor k para el cual se verifica que $P_k = 3,000$, siendo P_k el percentil de orden k . Dicho valor se encuentra en el intervalo 2,000 – 4,000; por tanto:

$$P_k = L_{i-1} + a_i \frac{\frac{Nk}{100} - N_{i-1}}{n_i}$$

$$3,000 = 2,000 + (4,000 - 2,000) \frac{\frac{193k}{100} - 0}{149}$$

$$1,000 = 25.9 k \quad k = 38.6$$

El 38.6% de los trabajadores casados cobran un salario de 3,000 pesos o menos y, en consecuencia, el 61.4 % de los trabajadores casados cobra más de 3,000 pesos.

- b. Se tiene que calcular el coeficiente de variación correspondiente a los trabajadores solteros y en unión libre.

$x_i = \text{salario}$	$n_i = \text{solteros}$	x'_i	$x'_i n_i$	$x_i'^2 n$
2,000 – 4,000	24	3,000	72,000	648,000,000,000
5,000 – 8,000	0	6,500	0	0
9,000 – 12,000	3	10,500	31,500	3,472,875,000,000
	27		103500	4,120,875,000,000

$$\bar{X}_s = \frac{\sum x'_i n_i}{N} = \frac{103500}{27} = 3833.3$$

$$S_s^2 = \frac{\sum x_i'^2 n_i}{N} - \bar{X}_s^2 = \frac{4,120,875,000,000}{27} - (3833.3)^2$$

$$= 152,610,306,811$$

$$S_s = 390,653.69$$

$$CV_s = \frac{S_s}{\bar{X}_s} = \frac{390653.69}{3833.3} = 101.91$$

$x_i = \text{salario}$	$n_i = \text{unión libre}$	x_i'	$x_i' n_i$	$x_i'^2 n$
2,000 – 4,000	21	3,000	63000	567,000,000,000
5,000 – 8,000	9	6,500	58500	2,471,625,000,000
9,000 – 12,000	0	10,500	0	0
	30		121500	3,038,625,000,000

$$\bar{X}_{ul} = \frac{\sum x_i' n_i}{N} = \frac{121500}{30} = 4050$$

$$S_{ul}^2 = \frac{\sum x_i'^2 n_i}{N} - \bar{X}_{ul}^2 = \frac{3,038,625,000,000}{30} - (4050)^2$$

$$= 101,271,000,000$$

$$S_{ul} = 318231.2$$

$$CV_{ul} = \frac{S_{ul}}{\bar{X}_{ul}} = \frac{318231.2}{4050} = 78.57$$

Como $CV_{ul} < CV_s$, se deduce que es más homogénea la distribución de salarios en los trabajadores en unión libre.

MEDIDAS DE POSICIÓN Y DISPERSIÓN

9. Utilice las columnas de información del año 2005 para calcular e interpretar los resultados de las medidas de tendencia central y las medidas de dispersión que se especifican a continuación:

	2005		2006	
Desc_Municipio	Volumen del consumo de energía eléctrica medido en Megawatts-hora	Usuarios de energía eléctrica	Volumen del consumo de energía eléctrica medido en Megawatts-hora	Usuarios de energía eléctrica
Acambay	28647	11927	15274	12235

Acolman	43902	10013	46626	9867
Aculco	74807	29557	25514	11043
Almoloya de Alquisiras	4734	5071	5018	5338
Almoloya de Juárez	171244	29791	180789	39556
Almoloya del Río	25927	1793	23115	1821
Amanalco	4406	4007	2832	2468
Amatepec	9793	9026	6248	6752
Amecameca	27374	11115	30083	11304
Apaxco	66593	4252	82656	4178
Atenco	9699	4098	9651	4040
Atizapán	2782	1477	2709	1497
Atizapán de Zaragoza	366010	107950	363726	108704
Atlacomulco	175549	24385	153568	25184

a) Medidas de tendencia central en el consumo de energía

Media, mediana, moda

Respuesta:

2005

Volumen del consumo de energía eléctrica medido en Megawatts-hora

Media	72247.64286
Error típico	27282.50711
Mediana	28010.5
Moda	#N/A
Desviación estándar	102081.7943
Varianza de la muestra	10420692719
Curtosis	4.910684278
Coefficiente de asimetría	2.166818698
Rango	363228
Mínimo	2782
Máximo	366010
Suma	1011467
Cuenta	14
Mayor (1)	366010
Menor(1)	2782
Nivel de confianza(95.0%)	58940.27323
Coefficiente de Variación	1.412942903

El promedio de consumo de energía eléctrica consumido en los diferentes municipios de México en el año 2005 fue de 72247.64286.

La mediana es decir el valor de los dos elementos intermedios de la muestra, ya que en este caso es una cantidad par es de 28010.5

La muestra no tiene ningún elemento que se repita más de una vez, por lo tanto no tiene Moda.

b) Medidas de Dispersión en la producción de energía

Rango, Varianza, Desviación Estándar y Coeficiente de variación

Respuesta:

En cuanto al Rango, la diferencia entre el dato mayor y el dato menor es de 363228.

La Desviación Estándar es decir la raíz cuadrada de la varianza es de 102081.7943.

La Varianza se define como el promedio de los cuadrados de las desviaciones de los datos con respecto a la media. Su valor indica la forma en que están distribuidos los datos con respecto a la media. Y con esta buscamos la confianza a las estimaciones de los datos, por lo tanto buscamos que tenga menor variabilidad, que este dato este muy cerca de la media y la desviación estándar, y en este caso la muestra presenta una alta variabilidad y su valor no se encuentra cercano al de la media y desviación estándar por lo tanto la confiabilidad de los datos es muy baja.

El Coeficiente de Variación en esta distribución de frecuencias es de 1.41%. Los datos presentan una alta variación.

10. Utilice los datos del año 2006 para determinar las medidas de posición que se especifican a continuación:

a) Medidas de posición en los usuarios de energía eléctrica

Percentil $\frac{1}{4}$, Percentil $\frac{3}{4}$, Desviación cuartílica

Respuesta:

Primer Cuartil	3.5	4074.5
Tercer Cuartil	10.5	12002.25

**Desviación
Cuartílica** **7927.75**

Si la desviación cuartílica es pequeña, significa que el 50% de las desviaciones se concentra en una zona pequeña y por lo tanto la dispersión es baja.

Por lo tanto la desviación cuartílica es 7927.75 es decir que las dispersiones en los datos es alta.

b) Medidas de asimetría en los usuarios de energía eléctrica

¿Se presenta simetría o asimetría en los datos?, Coeficiente de asimetría, Curtosis

Los datos están distribuidos de forma asimétrica . En una distribución simétrica tienen igual valor la media y la mediana, cuando es unimodal también coinciden con la moda. Y en esta distribución la media, mediana y moda no tienen igual valor.

El Coeficiente de Asimetría es de 2.99 por lo tanto esta distribución tiene asimetría positiva, con cola hacia la derecha.

El Coeficiente de Curtosis es de 9.58, entonces la curva de esta distribución es leptocúrtica o alargada.

11. Determine cuál es la variable dependiente e independiente para explicar el volumen de energía que se consume en el año 2006

Determine el coeficiente de Pearson para identificar la relación existente entre los usuarios y el volumen energético que demandan en el año 2006.

Respuesta:

El Volumen de consumo de energía depende de los usuarios de energía eléctrica que demanden de la misma.

Por lo tanto la variable dependiente es el Volumen de consumo de energía eléctrica y la variable independiente son los usuarios de energía eléctrica.

Covarianza de la muestra 2754947890

Coefficiente de Pearson 0.95357926

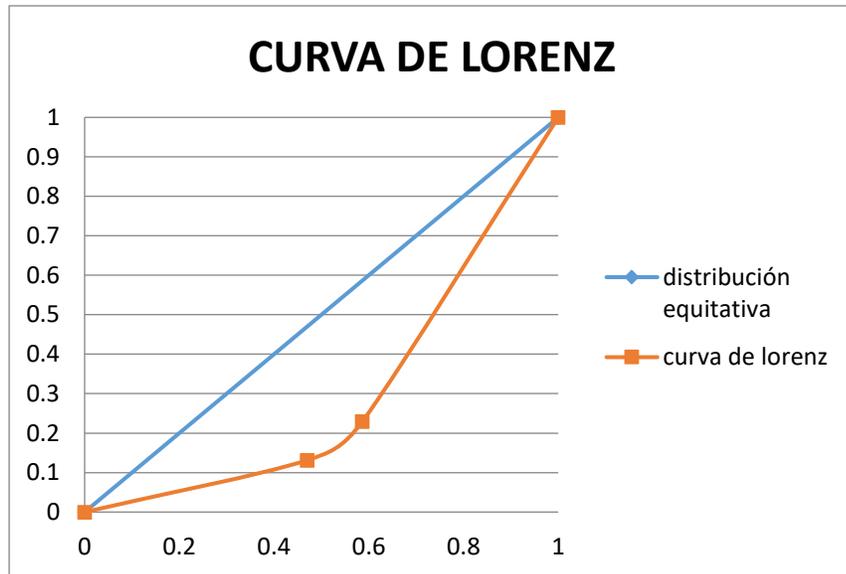
ANÁLISIS DE CONCENTRACIÓN

12. Construya e interprete la curva de Lorenz y el estadístico de Gini para los siguientes datos

Regiones Hidrológicas i	Superficie en KM cuadrados	Población para el 2012 en millones de habitantes
1	205218	2,7
2	145385	4,21
3	137753	4,34
4	152013	4,38
5	202562	4,42
6	77525	4,95
7	127166	5,14
8	101231	7,39
9	104790	10,31
10	119248	11,44
11	379552	11,84
12	16438	22,62
13	190367	23,29
Totales	1959248	117,03

Respuesta:

INTERVALOS	KM	POB MEDIA	POB GRPAL	%KM	% POB	CRECIENTES		DECRECIENTES	
						KM	POB	KM	POB
0 5000000	920456	2500000	2301140000000.00	0.47	0.13	0.47	0.13	1	1
5000001 10000000	228397	7500000.5	1712977614198.50	0.12	0.10	0.59	0.23	0.53	0.87
10000001 23290000	810395	16645000.5	13489025180197.50	0.41	0.77	1	1	0.41	0.77
	1959248	26645001	17503142794396.00	1	1				



La población para el año 2012 está distribuida de manera desigual en la superficie en KM cuadrados. El 10% de la población debería de estar ubicado en el 10% de la superficie y así sucesivamente para que la curva de Lorenz se distribuyera de manera equitativa.

ESTADÍSTICA INFERENCIAL

PRUEBA DE HIPÓTESIS

13. El Gobierno del Estado de México, sostiene que el agua sanitaria del municipio de Toluca es potable. Sin embargo, hay un grupo radical de oposición que asegura lo contrario. El agua se considera potable cuando el indicador de partículas en suspensión (IPS), normalmente tiene una media μ igual a 45 nano micras (nm) y una desviación estándar $\sigma = 2.5$. El Gobierno del Estado pide a los alumnos de primer semestre de la Maestría en Estudios Sustentables Regionales y Metropolitanos, para que estudien si el agua es o no potable. Como el IPS es desconocido y no se puede observar porque habría que analizar toda el agua sanitaria, los estudiantes con su amplio conocimiento de Estadística, específicamente en obtener muestras, analizan 15 observaciones de 1 mililitro tomadas en distintos puntos de la red de distribución. Para cada muestra se mide el indicador y se registra: 50, 46, 47, 45, 50, 53, 44, 40, 41, 45, 52, 55, 44, 48 y 45.

a) ¿Cuál es la probabilidad de que el IPS sea mayor a 46 nm?

Respuesta:

LA PROBABILIDAD ES DE 0.34457826.

b) ¿Se puede concluir con estos datos que el agua sanitaria es potable?

Respuesta:

NO ES POTABLE PORQUE Z ES MAYOR QUE 2. EL VALOR DE Z ES 2.76 Y PARA QUE SE PUDIERA CONCLUIR QUE SI LO ES TENDRIA QUE OSCILAR ENTRE -2 Y 2

c) Contrastar la H_0 de que esta muestra proviene de una población con media 45 y varianza σ^2 cualquiera, con un nivel de significancia $\alpha=5\%$, e intérprete; recuerde que $H_0: \mu = 45$ $H_1: \mu \neq 45$.

Respuesta:

ACEPTAMOS LA HIPOTESIS NULA DE QUE LA MUESTRA PROVIENE DE UNA POBLACIÓN CON MEDIA 45.

YA QUE EL VALOR T ENTRA EN LA REGIÓN CRÍTICA.

d) Como los alumnos de la MESRyM quieren entregar un análisis de pruebas de hipótesis bastante exhaustivo, deciden hacer otra muestra, obteniendo los siguientes registros: 52, 48, 44, 40, 57, 48, 54, 42, 39, 46, 54, 50, 49, 53 y 44.

¿Puede concluirse que la muestra extraída proviene de la misma red de distribución? (Haga la prueba de hipótesis para el caso de la igualdad de medias y varianzas con un nivel de significancia de $\alpha=10\%$)

Respuesta:

No, ya que rechazamos la hipótesis nula, porque el valor t no entra en la región crítica.

Es decir la muestra es diferente a la población, no forma parte de la población.

14. Para el siguiente ejercicio, según la Secretaria de Energía se cuenta con los siguientes registros:

PROMEDIO DE USUARIOS DE ENERGÍA ELÉCTRICA (miles)		
Servicios	Agrícola	Empresa Mediana
116	92	106
123	94	117
131	97	128
139	99	139
145	102	151
152	105	165
158	107	180
164	110	196
162	113	212
168	115	225
174	117	236
180	119	244
186	121	257
190	124	270

Que representan los usuarios por sector de productividad, excluyendo a la gran industria.

Con los datos anteriores, ¿Existe evidencia para considerar que los promedios de usuarios para los diferentes sectores productivos de México son iguales? (Considere un error de tipo II y un nivel de significancia de $\alpha = 10\%$).

Respuesta:

Análisis de varianza de un factor

RESUMEN

<i>Grupos</i>	<i>Cuenta</i>	<i>Suma</i>	<i>Promedio</i>	<i>Varianza</i>
Servicios	14	2188	156.285714	534.065934
Agrícola	14	1515	108.214286	108.027473
Empresa Mediana	14	2626	187.571429	2973.8022

ANÁLISIS DE VARIANZA

<i>Origen de las variaciones</i>	<i>Suma de cuadrados</i>	<i>Grados de libertad</i>	<i>Promedio de los cuadrados</i>	<i>F</i>	<i>Probabilidad</i>	<i>Valor crítico para F</i>
	44740.333			18.559855		2.4440438
Entre grupos	3	2	22370.1667	5	2.1705E-06	6
Dentro de los grupos	47006.642	9	1205.29853			
	91746.976					
Total	2	41				

Se rechaza la hipótesis nula.

Los promedios de usuarios de energía en los diferentes sectores productivos varían, según su producción y/o diferentes factores influyentes.

ANÁLISIS DE REGRESIÓN

15. Dada las siguientes series “índice nacional de precios al consumidor” (INPC) y “oferta monetaria” (M1), siga las siguientes indicaciones:

- obtenga el gráfico de las series
- Estime una regresión donde la variable dependiente es el INPC y evalúe si es una regresión espuria.

Series temporales del INPC y M1 para México.

Fecha	INPC	M1
Ene-90	15.01	27105212
Feb-90	15.35	27643193
Mar-90	15.06	28787121
Abr-90	15.86	30063052
May-90	16.14	32642282
Jun-90	16.49	33694120
Jul-90	16.79	33281548
Ago-90	17.08	33187651
Sep-90	17.32	33767717
Oct-90	17.57	38986751
Nov-90	18.04	43351826
Dic-90	18.60	50958966
Ene-91	19.08	46499678
Feb-91	19.41	47449447
Mar-91	19.69	49337176
Abr-91	19.89	50915227
May-91	20.09	53400232
Jun-91	20.30	54194748
Jul-91	20.48	52222535
Ago-91	20.62	56716695
Sep-91	20.83	64265971
Oct-91	21.07	92242680
Nov-91	21.59	99014690
Dic-90	22.10	113635550
Ene-92	22.50	106564446
Feb-92	22.77	104991259
Mar-92	23.00	102849442
Abr-92	23.21	104888362
May-92	23.36	109009455
Jun-92	23.52	109984259
Jul-92	23.67	109564288
Ago-92	23.81	107186849
Sep-92	24.02	105678924
Oct-92	24.19	116725201

Nov-92	24.39	120376602
Dic-92	24.74	131732775
Ene-93	25.05	125921361
Feb-93	25.25	126283697
Mar-93	25.40	124417159
Abr-93	25.55	124494574
May-93	25.69	128193168
Jun-93	25.84	130906015
Jul-93	25.96	132109404
Ago-93	26.10	130677665
Sep-93	26.30	131300992
Oct-93	26.40	134862001
Nov-93	26.52	137851027
Dic-93	26.72	157046723
Ene-94	26.93	151806822
Feb-94	27.07	150928567
Mar-94	27.21	149959703
Abr-94	27.34	143411158
May-94	27.47	144663410
Jun-94	27.61	145862730
Jul-94	27.73	146132380
Ago-94	27.86	143984532
Sep-94	28.06	144230158
Oct-94	28.21	146117260
Nov-94	28.36	152511658
Dic-94	28.61	163815904
Ene-95	29.68	144685439
Feb-95	30.94	138093959
Mar-95	32.76	128832393
Abr-95	35.37	125808463
May-95	36.85	126314945
Jun-95	38.02	128234014
Jul-95	38.80	133102217
Ago-95	39.44	132852879
Sep-95	40.26	135263950

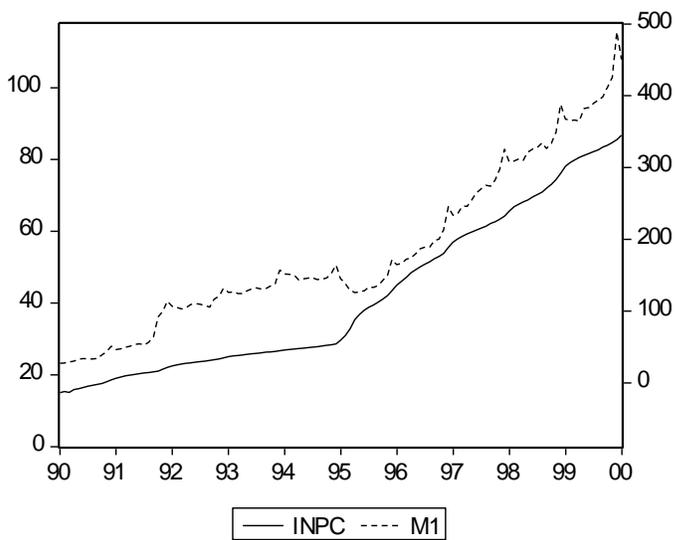
Oct-95	41.09	142980663
Nov-95	42.10	149977213
Dic-95	43.47	171641773
Ene-96	45.03	164646328
Feb-96	46.08	166415710
Mar-96	47.10	172326184
Abr-96	48.44	174278059
May-96	49.32	179958385
Jun-96	50.12	186763642
Jul-96	50.84	188996330
Ago-96	51.51	189094573
Sep-96	52.34	197632138
Oct-96	52.99	200322278
Nov-96	53.79	212751512
Dic-96	55.51	245258434
Ene-97	56.94	233232562
Feb-97	57.90	235616681
Mar-97	58.62	246307982
Abr-97	59.25	245721391
May-97	59.79	254640695
Jun-97	60.32	264505195
Jul-97	60.85	270367093
Ago-97	61.39	275211063
Sep-97	62.15	274028115
Oct-97	62.65	282170648
Nov-97	63.35	298075076
Dic-97	64.24	325389290
Ene-98	65.64	308002734
Feb-98	66.79	308796065
Mar-98	67.57	312067625
Abr-98	68.20	310030942
May-98	68.74	321651818
Jun-98	69.56	325901070
Jul-98	70.23	328168113
Ago-98	70.90	334274707

Sep-98	72.05	326337621
Oct-98	73.09	333589954
Nov-98	74.38	348787013
Dic-98	76.19	387895562
Ene-99	78.12	367402964
Feb-99	79.17	365178625
Mar-99	79.90	366012077
Abr-99	80.64	364552522
May-99	81.12	382397418
Jun-99	81.66	383631018
Jul-99	82.20	389526912
Ago-99	82.66	393472895
Sep-99	83.46	398432301
Oct-99	83.99	411617056
Nov-99	84.73	425224824
Dic-99	85.58	489132958
Ene-00	86.73	450678580

Fuente: INEGI

Respuesta:

a)



b) Si tomamos a la variable INPC (Y) como variable dependiente y a la variable M1 (X) como variable independiente. Supóngase que se efectúa la regresión de INPC sobre M1, utilizando la información de la tabla 1, los resultados son los siguientes:

$$INPC = 7.040392 + 0.190014M1$$

$$t = (7.7673) \quad (45.1674)$$

$$R^2 = 0.9448 \quad d = 0.1639$$

Los resultados que se obtienen de esta regresión se ven muy bien ya que podemos observar un R^2 bastante alto (0.9448), al igual que la razón t de M1 (45.1674). El único resultado que no se ve del todo bien es el Durbin- Watson ya que d es bajo (0.1639). Granger y Newbold, han sugerido una regla práctica para sospechar que la regresión estimada sufre de regresión espuria es: $R^2 > d$ (0.9448 > 0.1639). Por tal motivo podemos sospechar que nuestra regresión es espuria.

BIBLIOGRAFÍA

Box., G. William H. y Hunter.S. (2008) *Estadística para investigadores*. México. Editorial Reventé S. A.

César Pérez (2009) Técnicas de Análisis de Datos con SPSS 15. Biblioteca Virtual Pearson
<https://www.biblionline.pearson.com/Pages/BookDetail.aspx?b=442>

David M. Lane (2013) Introduction to Statistics. Autoedición 692 pág.
<https://openlibra.com/es/book/introduction-to-statistics>

García Pérez, A. (2015) La interpretación de los datos. Una introducción a la Estadística Aplicada. UNED. Madrid, España

Llinás, S., H. y Rojas, A.C. (2015) Estadística Descriptiva y distribuciones de probabilidad. Barranquilla Colombia: Universidad de Norte.

Mayra L. Samuels (2012) Fundamentos de Estadística para las Ciencias de la Vida. Biblioteca Virtual Pearson
<https://www.biblionline.pearson.com/Pages/BookDetail.aspx?b=929>

Peña, Daniel (2014) Fundamentos de Estadística. Alianza Editorial. Madrid.