

Nombre de la alumna

Andry Pracseli Castro Hernández

Nombre del Catedrático

Rosario Gómez Lujano

Nombre del trabajo

Ensayo

Materia

Estadística I

Grado

1° Cautrimestre

Grupo

Unico

Licenciatura

Psicología

Pichucalco, Chis. A 26 de septiembre de 2000

Estadística

"Estadística" se derivó de la palabra "Estado" la función de los gobiernos entre otras cosas es llevar los registros de población, nacimientos, cosechas, impuestos y toda la información que engloba el estado, es así que, tradicionalmente se definió a la estadística como un instrumento de compilación, organización, presentación y análisis de datos numéricos.

La estadística es una ciencia aplicada de las matemáticas y es una valiosa herramienta para la toma de decisiones. Permite el estudio de fenómenos mediante la descripción del mismo a través de inferencias mediante distribuciones probabilísticas. La estadística se ocupa de los métodos y procedimientos para recoger, clasificar, resumir, hallar regularidades y analizar los datos, siempre y cuando la variabilidad e incertidumbre sea una causa intrínseca de los mismos; así como de realizar inferencias a partir de ellos, con la finalidad de ayudar a la toma de decisiones y en su caso formular predicciones. La estadística fue fundada por el londinense Jonh Graunt "un mercader de mercería", en un pequeño libro "Natural and political observations made upon the bells of mortality". Este libro fue el primer intento para interpretar fenómenos biológicos de masa y de la conducta social.

La estadística es de gran importancia en las diferentes empresas, enfocadas desde cualquier área profesional ya que ayudan a lograr una adecuada planeación y control apoyados en los estudios de pronósticos presupuestos, etc.

-Clasificación de la estadística

La estadística se clásica en dos: estadística descriptiva y estadística inferencial.

● **Estadística descriptiva**: se sustituye o reduce el conjunto de datos obtenidos por un pequeño número de valores descriptivos, como pueden ser: el promedio, la mediana, la media geométrica, la varianza, la desviación típica, etc. Éstas medidas descriptivas pueden ayudar a brindar las principales propiedades de los datos observados así como las características clave de los fenómenos bajo investigación.

Por lo general la información proporcionada por la estadística descriptiva puede ser transmitida con facilidad y eficacia mediante una variedad de herramientas gráficas, como puede ser:
Gráficos de tendencia: es un trazo de una característica de interés sobre un período para observar su comportamiento en el tiempo.

Gráfico de dispersión :ayuda al análisis de la relación entre dos variables, representado gráficamente sobre el eje X y el correspondiente valor de la otra sobre el eje Y.

Histograma:describe la distribución de los valores de una característica de interés.

Describe, analiza y representa un grupo de datos utilizando métodos numéricos y gráficos que resumen y presentan la información contenida en ellas. La estadística descriptiva es aplicable en casi todas las áreas donde se recopilan datos cuantitativos. Puede brindar información acerca de productos, procesos o diversos aspectos del sistema de gestión de la calidad, como también en el hábito de la dirección y organización de personas, la logística, etc.

● **Estadística inferencial:** la estadística inferencial emplea usualmente mecanismos que le permiten llevar a cabo dichas deducciones, tales como pruebas de hipótesis, pruebas paramétricas (como de medida, de diferencia de medias, proporciones, etc.) y no paramétricas. La estadística inferencial es sumamente útil en el análisis de poblaciones y tendencias, para hacerse una idea posible de las acciones y reacciones de la misma de cara a condiciones específicos. Esto no significa que se las pueda predecir fielmente, ni que estemos en presencia de una ciencia exacta, pero si de una aproximación posible al resultado final.

-Población

Es total de individuos o conjunto de ellos que presentan o podrían presentar el rasgo característico que se desea estudiar. Una población puede ser finita o infinita

● **Población finita:** es aquella en la que el número valores que la componen tiene un fin. Por ejemplo, la población estadística que nos indica la cantidad de árboles de una ciudad en finita. Es cierto que puede variar con el tiempo, pero en un instante determinado es finita, tiene fin.

● **Población infinita:** se trata de aquella población que no tiene fin. La que tiene un elemento incontable de elementos. Son aquellas de tamaño desconocido.

-Muestra

consiste en la porción que se extrae de una población estadística para un determinado estudio, con el fin de representar, conocer y determinar los aspectos de dicha población.

Este método se utiliza cuando no se puede aplicar un censo en ciertas poblaciones, ya que a través del muestreo se puede establecer una porción de la realidad a estudiar.

Tipos de muestra estadística:

La muestra estadística puede clasificarse en muestra probabilística y no probabilística.

El muestreo probabilístico es el tipo más utilizado durante las investigaciones. En este, todos los elementos de la población o universo tienen la posibilidad de ser parte de la muestra.

Dentro de la muestra probabilística podemos encontrar los siguientes tipos de muestreo:

● **Muestreo aleatorio simple:** método de selección más básico, en el que cada sujeto posee un número de identificación y, mediante un sorteo aleatorio, algunos son seleccionados para la muestra. Para realizarla se debe conocer la cantidad de individuos que serán necesarios para completar toda la muestra.

● **Muestreo sistemático:** con este método se enumera la población a trabajar y los investigadores se encargan de listar a cada individuo en grupos de 10, de forma aleatoria. Luego se elige a uno de los primeros de cada grupo al azar, formando de esta manera la muestra.

● **Muestreo estratificado:** consiste en dividir a la población en estratos o grupos que compartan características similares. Luego se seleccionan proporcionalmente algunos individuos de cada grupo.

● **Muestreo por conglomerados:** se da cuando la población ya se encuentra dividida de forma natural en grupos, por lo que se seleccionan individuos aleatoriamente de cada conglomerado para conformar la muestra.

Muestra no probabilística

En la muestra no probabilística, los elementos se seleccionan a través de procesos que no brindan, a todos los individuos de la población, la misma posibilidad de ser elegidos para la muestra.

A su vez, este deriva en los siguientes tipos de muestreo:

- *Muestreo por cuotas*: los investigadores forman la muestra partiendo de determinadas características, con el fin de lograr la misma distribución de características en la muestra que en la población. Por ejemplo, si en una población hay 70% hombres y 30% mujeres, los individuos de la muestra deberán estar distribuidos de la misma manera.

- *Muestreo por conveniencia*: aquí el investigador suele elegir a los individuos de su muestra solo por su proximidad. Generalmente, el investigador no reconoce esta muestra como una representación de toda población, sino que la realiza para conocer opiniones, datos e información de manera rápida.

- *Muestreo por bola de nieve*: se utiliza cuando el investigador requiere que un sujeto de su muestra ayude a identificar a otro con las mismas características y, a su vez, estos a otros, para formar la muestra deseada.

- *Muestra discrecional*: también conocida como muestreo por juicio o criterio, se da cuando el investigador selecciona los individuos de su muestra en base a determinado conocimiento de la población. Por ejemplo, si desea realizar una investigación de personas con problemas psicológicos en el embarazo, seleccionará aquellas mujeres que hayan pasado por la experiencia del embarazo.

-Muestreo

En ocasiones en que no es posible o conveniente realizar un censo (analizar a todos los elementos de una población), se selecciona una muestra, entendiendo por tal una parte representativa de la población.

El muestreo es por lo tanto una herramienta de la investigación científica, cuya función básica es determinar que parte de una población debe examinarse, con la finalidad de hacer inferencias sobre dicha población.

La muestra debe lograr una representación adecuada de la población, en la que se reproduzca de la mejor manera los rasgos esenciales de dicha población que son importantes para la investigación. Para que una muestra sea representativa y por lo tanto sea útil, debe reflejar las similitudes y diferencias encontradas en la población, es decir ejemplificar las características de esta.

El muestreo es imprescindible en la investigación, ya sea ésta de cualquier ciencia aplicada, también ha sido el proceso por el medio del cual algunas disciplinas han podido introducir en ellas metodologías y procedimientos para su consolidación como tal, una de las grandes disciplinas beneficiadas en la administración.

-Censo

Del latín census, un censo es un padrón o lista. Su uso más habitual está asociado al censo poblacional o censo de población, donde se contabilizan los habitantes de un pueblo o de un país y se recogen diversos datos con fines estadísticos.

Por ejemplo: “El Gobierno anunció que el año próximo realizará un censo nacional”, “Los datos recopilados a través del censo permitirán desarrollar políticas sociales más precisas”, “De acuerdo al último censo, la ciudad ya ha superado los dos millones de habitantes”.

El censo permite delimitar una población estadística que refleja el número total de individuos de un territorio. A diferencia de otras herramientas de la estadística, en este caso no se trabaja con muestras, sino con la población total. El censo es considerado una técnica que utiliza la ficha censal o cédula de censo como instrumento para llevar a cabo una investigación.

-Variables

Variable es una característica (magnitud, vector o número) que puede ser medida, adoptando diferentes valores en cada uno de los casos de un estudio.

Variables cualitativas: Son las variables que expresan distintas cualidades, características o modalidad. Cada modalidad que se presenta se denomina atributo o categoría y la medición consiste en una clasificación de dichos atributos. Las variables cualitativas pueden ser ordinales y nominales. Las variables cualitativas pueden ser dicotómicas cuando sólo pueden tomar dos valores posibles como sí y no, hombre y mujer o son politómicas cuando pueden adquirir tres o más valores.

Dentro de ellas podemos distinguir:

■ **Variable cualitativa ordinal:** La variable puede tomar distintos valores ordenados siguiendo una escala establecida, aunque no es necesario que el intervalo entre mediciones sea uniforme, por ejemplo, leve, moderado, grave.

■ **Variable cualitativa nominal:** En esta variable los valores no pueden ser sometidos a un criterio de orden como por ejemplo los colores o el lugar de residencia.

Variables cuantitativas: Son las variables que se expresan mediante cantidades numéricas. Las variables cuantitativas además pueden ser:

● **Variable discreta:** Es la variable que presenta separaciones o interrupciones en la escala de valores que puede tomar. Estas separaciones o interrupciones indican la ausencia de valores entre los distintos valores específicos que la variable pueda asumir.

Ejemplo: El número de hijos (1, 2, 3, 4, 5).

● **Variable continua:** Es la variable que puede adquirir cualquier valor dentro de un intervalo especificado de valores. Por ejemplo el peso (2,3 kg, 2,4 kg, 2,5 kg, ...) o la altura (1,64 m, 1,65 m, 1,66 m, ...), que solamente está limitado por la precisión del aparato medidor, en teoría permiten que siempre exista un valor entre dos cualesquiera.

-Medidas de tendencia central

Se les llama medidas de tendencia central a la media aritmética, la mediana, la media geométrica, la moda, etc. debido a que al observar la distribución de los datos, éstas tienden a estar localizados generalmente en su parte central

Media aritmética: también se le conoce como promedio ya que es el promedio de las lecturas o mediciones individuales que se tienen en la muestra.

Mediana: es aquel valor que se encuentra en la parte central de los datos que se tienen en la muestra una vez que estos han sido ordenados según su valor o magnitud.

Media geométrica: es la raíz en n -ésima del producto de los valores de los elementos de la muestra, es usada cuando los valores de los datos de la muestra no son lineales, es decir que su valor depende de varios factores a la vez.

Moda: la moda se define como aquel valor o valores que más se repiten o que tienen mayor frecuencia entre los datos que se han obtenido en una muestra.

-Medidas de dispersión

Cuando se tiene una muestra de datos obtenida de una población cualquiera, es importante determinar sus medidas de tendencia central así como también es básico el determinar que tan dispersos están los datos en la muestra, por lo que se hace necesario determinar si rango, la varianza, la desviación estándar, etc., ya que una excesiva variabilidad o dispersión en los datos indica la inestabilidad del proceso en análisis en la mayoría de los casos.

Ejemplos de distribución de frecuencia no agrupados

Ejemplo 1:

Se le pidió a un grupo de personas que indiquen su color favorito, y se obtuvo los siguientes resultados:

Negro	Azul	Amarillo	Rojo	Azul
Azul	Rojo	Negro	Amarillo	Rojo
Rojo	Amarillo	Amarillo	Azul	Rojo
Negro	Azul	Rojo	Negro	Amarillo

Con los resultados obtenidos, elaborar una tabla de frecuencias

Solución:

En la primera columna, colocamos los valores de nuestra variable, en la segunda la frecuencia absoluta luego la frecuencia relativa y finalmente la frecuencia relativa acumulada

Color	Frecuencia absoluta	Frecuencia acumulada	Frecuencia relativa	Frecuencia relativa acumulada
Negro	4	4	0,20	0,20
Azul	5	9	0,25	0,45
Amarillo	5	14	0,25	0,70
Rojo	6	20	0,30	1
Total	20		1	

Ejemplo 2:

En una tienda de autos, se registra la cantidad de autos Toyota vendidos en cada día del mes de septiembre.

0,1,2,1,2,0,3,2,4,0,4,2,1,0,3,0,0,3,4,2,0,1,1,3,0,1,2,1,2,3

Con los datos obtenidos, elaborar una tabla de frecuencias

Solución

En la primera columna colocamos los valores de nuestra variable, en la segunda la frecuencia acumulada, seguida por la frecuencia relativa acumulada. Ahora vamos a agregar la columna de frecuencia porcentual y frecuencia porcentual acumulada

Autos vendidos	Frecuencia absoluta	Frecuencia acumulada	Frecuencia relativa	Frecuencia relativa acumulada	Frecuencia porcentual	Frecuencia porcentual acumulada
0	8	8	0,267	0,267	26,7%	26,7%
1	7	15	0,233	0,500	23,3%	50,0%
2	7	22	0,233	0,733	23,3%	73,3%
3	5	27	0,167	0,900	16,7%	90,0%
4	3	30	0,100	1	10,0%	100%
Total	30		1		100%	

Ejemplo de distribución de frecuencia agrupados

Ejemplo 1

Las notas de 35 alumnos en el examen final de estadística, calificado del 0 al 10, son las siguientes:

0,0,0,0,1,1,1,1,2, 2,2,3,3,3,3,4,4,4,4,5,5,5,5,6,6,6,7,7,7,8,8,8,9,10,10.

Con los datos obtenidos, elaborar una tabla de frecuencias con 5 intervalos o clases.

Solución:

-Hallamos el rango: $R = X_{max} - X_{min} = 10 - 0 = 10$

-El número de intervalos (k), me lo da el enunciado del programa: $k = 5$

-Calculamos la amplitud de clase: $A = R / K = 10 / 5 = 2$

-Ahora hallamos los límites inferiores y superiores de cada clase y elaboraremos la tabla de frecuencia.

Intervalo	Marca de clase	Frecuencia absoluta	Frecuencia acumulada	Frecuencia relativa	Frecuencia relativa acumulada
0-2	1	8	8	0,229	0,229
2-4	3	7	15	0,200	0,429
4-6	5	8	23	0,229	0,658
6-8	6	6	29	0,171	0,829
8-10	9	6	35	0,171	1
Total		35		1	

Ejercicio 2

Un grupo de atletas se está preparando para una maratón siguiendo una dieta muy estricta. A continuación, viene el peso en kilogramos que ha logrado bajar cada atleta gracias a la dieta y ejercicios

0,2	8,4	14,3	6,5	3,4
4,6	9,1	4,3	3,5	1,5
6,4	15,2	16,1	19,8	5,4
12,1	9,6	8,7	12,1	3,2

Elaborar una tabla de frecuencias con dichos valores.

Solución

-Hallamos el rango : $R = X_{\max} - X_{\min} = 19,8 - 0,2 = 19,6$.

-El número de intervalos (K), lo calculamos usando la regla de sturges: $k = 1 + 3,322 \log(n) = 1 + 3,322 \log(20) = 5,32$. Podemos redondear el valor de k a 5

-Calculamos la amplitud de clase: $A = R/K = 19,6/5 = 3,92$ redondeados a 4

Ahora hallamos los límites inferiores y superiores de cada clase y elaboramos la tabla de frecuencias

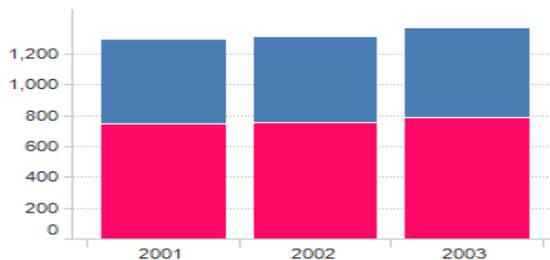
Intervalo	Marca de clase	Frecuencia absoluta	Frecuencia acumulada	Frecuencia relativa	Frecuencia relativa acumulada
0-4	2	5	5	0,25	0,25
4-8	6	5	10	0,25	0,50
8-12	10	4	14	0,20	0,70
12-16	14	4	18	0,20	0,90
16-20	18	2	20	0,10	1
Total		20		1	

Ejemplo y definición de gráfica de barras, de pastel o circular, histogramas, polígono de frecuencia.

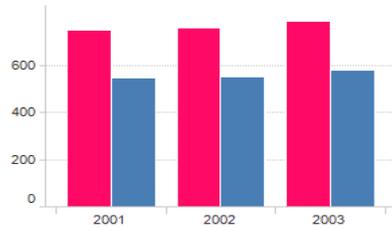
Gráfica de barra: Los gráficos de barra son aquellos que revelan cotejos entre elementos individuales. En ocasiones este tipo de gráficas, las categorías se muestran organizadas de manera vertical; mientras que los valores se ordenan horizontalmente.

Ejemplo

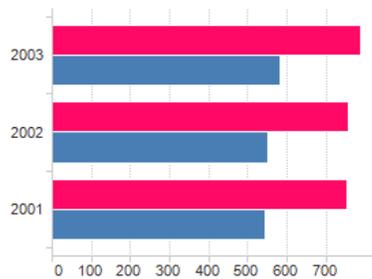
Una tabla de datos contiene las cifras de ventas de varias frutas y verduras distintas. El gráfico de barra puede mostrar la suma total de ventas de varios años.



Las barras están coloreadas según la columna categoría, que contiene los dos valores: frutas y verduras.



En la segunda imagen, se ha agregado la columna categoría al eje X, de manera que las barras sean divididas en dos barras distintas, una para cada categoría.



Las barras también pueden mostrarse de forma horizontal

Gráfica de pastel o circular: El gráfico circular es aquel que indica el tamaño proporcional de los elementos que componen una serie de datos basándose en la suma de sus elementos. Como resultado, debe mostrar una única serie de datos.

Histograma: describe la distribución de valores de una característica de interés. Estos métodos gráficos son de mucha utilidad para entender con claridad un fenómeno analizado. La evolución de la inflamación, el tipo de cambio, del PBI u otros indicadores macro pueden ser analizados, por ejemplo, con gráficos de tendencia.

Polígono de frecuencia: Polígono de frecuencia es el nombre que recibe una clase de gráfico que se crea a partir de un histograma de frecuencia. Estos histogramas emplean columnas verticales para reflejar frecuencias: el polígono de frecuencia es realizado uniéndolos puntos de mayor altura de estas columnas.



*Calcula la media aritmética o promedio, moda, rango, desviación estándar, varianza para los siguientes calificaciones que obtuvo un alumno LPSI de la UDS: 6,7,8,9,9,8,7,6,10

Media aritmética:

$$X = \frac{6+7+8+9+9+8+7+6+10}{9} = 7.7$$

Moda:

Promedio	Frecuencia
6	2
7	2
8	2
9	2
10	1

Rango:

$$R = VM - Vm$$

$$VM = 10. \quad Vm = 6$$

$$R = 10 - 6 = 4$$

Varianza:

$$\frac{2.89+0.49+0.09+1.69+1.69+0.09+0.49+2.89+5.29}{8} = \frac{15.61}{8} = 1.951$$

Desviación estándar:

$$= 1.3967$$