

INFORMÁTICA APLICADA A LA ESTADÍSTICA

SEMANA 1.

Temas	Enlaces Recomendados
Generalidades de la estadística	http://www.slideshare.net/produceideas/estadistica-descriptiva-presentation http://www.slideshare.net/hectorquintero/conceptos-bsicos-de-la-estadstica
La estadística descriptiva	
Concepto de base de datos	
Entorno gráfico de Excel	http://www.youtube.com/watch?v=L6oFoc-uxSw
Entorno gráfico de Statgraphics	

GENERALIDADES DE LA ESTADÍSTICA

Como sabemos la estadística define como la ciencia que recopila, organiza, analiza e interpreta la información numérica o cualitativa, mejor conocida como datos, de manera que se pueda llevar a conclusiones válidas. Dentro de la estadística se presenta el término dato, la cual es la base para trabajar en estadística, estos pueden ser cualitativos y cuantitativos. A los datos cuantitativos y cualitativos se los relaciona con el término variable, que es una característica susceptible de ser medida en las unidades de análisis que se estudian, que toma diferentes valores o grados de intensidad, en dependencia de cuál sea la unidad medida.

VARIABLES CUALITATIVAS O ATRIBUTOS		VARIABLES CUANTITATIVA	
NOMINALES	ORDINALES	DISCRETAS	CONTINUAS
Son números o letras que representan categorías donde no interesa el orden. Por ejemplo: 0 = masculino 1 = femenino	Son números o letras que representan categorías donde el orden interesa. Por ejemplo: 1 = Totalmente de acuerdo 2 = De acuerdo 3 = NS/NR 4 = Desacuerdo 5 = Totalmente desacuerdo	Son aquellos que surgen por el procedimiento de conteo. Es decir, los datos discretos toman valores enteros. Por ejemplo: número de hijos por familia, número de pacientes atendidos por un hospital.	Son aquellos que surgen cuando se mide alguna característica. Es decir, toman al menos teóricamente cualquier valor dentro de un intervalo. Por ejemplo: el peso, la estatura.

LA ESTADÍSTICA DESCRIPTIVA

El curso se centrará en la estadística descriptiva la cual recopila, organiza e interpreta la información numérica ó cualitativa. Un ejemplo de esta son los periódicos, las revistas, la radio y la televisión, ya que usan la estadística descriptiva para informar y persuadirnos acerca de ciertas acciones a tomar y en la formación de opiniones.

Cuando se trabaja en estadística es necesario tener en cuenta conceptos como universo o población y muestra. La población hace referencia a la colección de todos los individuos, objetos u observaciones que poseen al menos una característica en común, por ejemplo las edades de los estudiantes de la Universidad de Nariño. La muestra es una parte o subconjunto representativo de la población y al proceso de obtener una muestra se llama muestreo. Siguiendo el ejemplo anterior la muestra podría ser la edad de los estudiantes de primer semestre de los diferentes programas de la Universidad de Nariño.



TIPOS DE TABLAS.

Cuando se tiene cierto numero de datos, es necesario tener en claro que la tabulación de los mismos debe darse en forma ordenada. Por ejemplo no es lo mismo tener 5 datos que tener cien o mil datos. Por ello existen tres tipos de tablas.

* Tablas Tipo I: Cuando el tamaño de la muestra y el recorrido de la variable son pequeños, por ejemplo si tenemos una muestra de las edades de 5 personas, por lo que no hay que hacer nada especial simplemente anotarlas de manera ordenada en filas o columnas.

Ejemplo: Edad de los 5 miembros de una familia:

5, 8, 16, 38, 45

* Tablas Tipo II: Cuando el tamaño de la muestra es grande y el recorrido de la variable es pequeño, por lo que hay valores de la variable que se repiten. Por ejemplo, si preguntamos el número de personas activas que hay en 50 familias obtenemos la siguiente tabla:

Ejemplo:

Personas activas en 50 familias

2	1	2	2	1	2	4	2	1	1
2	3	2	1	1	1	3	4	2	2
2	2	1	2	1	1	1	3	2	2
3	2	3	1	2	4	2	1	4	1
1	3	4	3	2	2	2	1	3	3

Podemos observar que la variable toma valores comprendidos entre 1 y 4, por lo que resumamos estos datos quedando la siguiente tabla:

Personas activas	Número de familias
1	16
2	20
3	9
4	5
Total	50

* Tablas Tipo III: Cuando el tamaño de la muestra y el recorrido de la variable son grandes, por lo que será necesario agrupar en intervalos los valores de la variable. Por ejemplo si a un grupo de 30 alumnos les preguntamos el dinero que en ese momento llevan encima, nos encontramos con los siguientes datos:

450 1150 250 300 200 100 50 2600 600 750 1550 2300 5000
1200 100 50 200 650 500 400 1500 200 950 200 150 300
400 550 1100 150

Evidentemente, la variable estadística tiene un recorrido muy grande, 5000 pesos, por lo que si queremos hacer una tabla con estos datos tendremos que tomar intervalos. Para decidir la amplitud de los intervalos, necesitamos decidir ¿Cuántos intervalos queremos?. Normalmente se suele trabajar con no mas de 10 a 12 intervalos.

$\text{Amplitud} = 5000 / 10 = 500$ por lo que tomaremos intervalos de amplitud de 500

Debemos tener en cuenta las siguientes consideraciones: Tomar pocos intervalos implica que la pérdida de información sea mayor. Los intervalos serán siempre Cerrados por la izquierda y Abiertos por la Derecha [$Li-1$, Li). Procuraremos que en la decisión de intervalos los valores observados no coincidan con los valores de los extremos del intervalo y si esto ocurre que no sea en más de un 5% del total de observaciones.

Con estas recomendaciones tendremos la siguiente tabla:

[$Li-1$, Li)	Frecuencia
[0,500)	16
[500,1000)	6
[1000,1500)	3

[1500,2000)	2
[2000,2500)	1
[2500,3000)	1
[3000,3500)	0
[3500,4000)	0
[4000,4500)	0
[4500,5000)	1

CONCEPTO DE BASE DE DATOS

En estadística, un dato son los valores cualitativos o cuantitativos mediante los cuales se miden las características de los objetos, sucesos o fenómenos a estudiar

Una base de datos es un “almacén” que nos permite guardar grandes cantidades de información de forma organizada para que luego podamos encontrar y utilizar fácilmente. Desde el punto de vista informático, la base de datos es un sistema formado por un conjunto de datos almacenados en discos que permiten el acceso directo a ellos y un conjunto de programas que manipulen ese conjunto de datos.

Cada base de datos se compone de una o más tablas que guarda un conjunto de datos. Cada tabla tiene una o más columnas y filas. Las columnas guardan una parte de la información sobre cada elemento que queramos guardar en la tabla, cada fila de la tabla conforma un registro.

ENTORNO GRÁFICO DE EXCEL

ENTORNO GRÁFICO DE STATGRAPHICS

Esta parte de la primera semana se encuentra en el archivo ENTORNO GRÁFICO

NETGRAFÍA

Concepto de estadística, disponible en <http://ponce.inter.edu/cremc/estadistica.htm>

Concepto de variable, disponible en:

http://www.hospitalameijeiras.sld.cu/hha/images/presentaciones/investigaciones/Bayarre_y_Hosford.pdf

<http://www.mat.uda.cl/hgomez/Apuntes/Estad%C3%ADstica%20Descriptiva%20I.pdf>

Tipos de Tablas, disponible en: <http://thales.cica.es/rd/Recursos/rd97/UnidadesDidacticas/53-1-u-punt12.html>

Concepto de base de datos, disponible en;

<http://www.maestrosdelweb.com/principiantes/%C2%BFque-son-las-bases-de-datos/>