Nombre del alumno: Lesly Merari Utrilla López.

Nombre del profesor: Rosario Gómez Lujano.

Materia: Estadística 1.

Trabajo: Ensayo.

Grado: Primer semestre.

Grupo: Escolarizado psicología.

Pichucalco, Chiapas a 15 de noviembre del 2020.

I. Realizar un ensayo con los conceptos requeridos.

"La estadística y sus conceptos"

La estadística es "la disciplina que intenta sacar conclusiones de estudios empíricos mediante la utilización de modelos matemáticos. Sirve como nexo entre los fenómenos reales y los modelos matemáticos", en otras palabras, es una materia que se caracteriza por hacer uso de diversos procesos para logar un objetivo específico. Es una gran ayuda en nuestra vida cotidiana, por medio de ella resolvemos problemáticas y cuestiones que sean dificultosas. Cuenta con dos clasificaciones que son: Descriptiva e inferencial. Por ende, necesitamos comprender y entender su importancia, todos los procesos y conceptos que esta engloba. Tomando en cuenta que es muy completa y ocasionalmente tediosa.

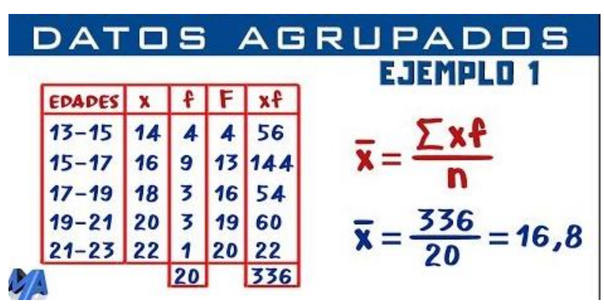
En esta ocasión abordaremos a cerca de las medidas de tendencia central, medidas de variabilidad para datos agrupados y no agrupados, sobre los cuantiles y coeficientes de variación y de asimetría.

Primero debemos tener en cuenta que las medidas de tendencia central "las medidas de tendencia central son medidas estadísticas que pretenden resumir en un solo valor a un conjunto de valores. Representan un centro en torno al cual se encuentra ubicado el conjunto de los datos". Las más empleadas son media, mediana y moda.

La media, se conoce con los nombres de media aritmética o promedio aritmético, no es más que un cálculo en que se promedian todos los datos obtenidos en la muestra. Para obtenerla necesitamos sumar todas las cantidades y dividirlas entre todos los datos. Es la más utilizada tradicionalmente en las ciencias porque es muy eficiente. Capta la variación de cada una de las puntuaciones. La media aritmética de la muestra se representa con x y guion arriba (-).

Ejemplo:

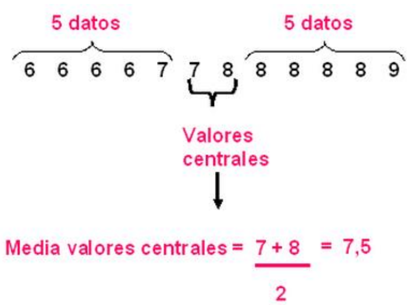
$$\bar{X} = \frac{8+9+10+11+16+17+6}{7} = \frac{77}{7} = 11$$



La media es de valores que más funciones cotidianas tiene, esto debido a que es responsable de arrojar un valor especifico. Es indispensable en casi todos los campos de estudio. Es un arma indispensable para profesores, estudiantes, investigadores, mercaderes, entre muchos otros. Es de vital importancia para los seres humanos.

La mediana es "el valor de la variable que ocupa la posición central, cuando los datos se disponen en orden de magnitud. Es decir, el 50% de las observaciones tiene valores iguales o inferiores a la mediana y el otro 50% tiene valores iguales o superiores a la mediana". Si se nos presentara el caso en el que contáramos con una cantidad de valores par, la mediana correspondiente será igual al promedio de los valores centrales. Si queremos obtener la posición de la mediana tenemos que tomar el número de datos, sumarle uno y dividirlo entre el dos. Es un estadístico máximamente eficiente y mínimamente robusto. No es un buen estadístico de tendencia central cuando en la distribución hay valores extremos o cuando es muy asimétrica. Cada conjunto tiene una y sólo una media. Para datos agrupados, la mediana se encuentra en el intervalo donde la frecuencia acumulada llega hasta la mitad de la suma de las frecuencias absolutas.

Ejemplos:



	Intervalos	Marca de clase x	Frecuencia absoluta f	Frecuencia acumulada F
1	[0 - 4)	2	3	3
2	[4 - 8)	6	5	8
3	**[8 - 12)**	**10**	**6**	**14**
4	[12 - 16)	14	4	18
5	[16 - 20)	18	3	21
	Total		21	

Por último, la moda "se define como el valor de la variable que más se repite. En un polígono de frecuencia la moda corresponde al valor de la variable que está bajo el punto más alto del gráfico. Una muestra puede tener más de una moda". Es el valor que representa la mayor frecuencia absoluta. En tablas de frecuencias con datos agrupados, hablaremos de intervalo modal. Para este solo requerimos de una buena observación. Pueden existir series de datos unimodales que son los más comunes, los bimodales, etc., siempre y cuando se repitan la misma cantidad de veces.

Ejemplos:

N° de Calzado	frecuencia
36	8
37	12
38	9
39	7
40	6
41	2
42	2
Total	46

- 3, 5,3,8,5,3,6,8,1,3,9,4,3,6,8

La moda es 3

- 1,6,1,5,3,8,1,5,7,5,1,2,3,5

La moda es 1 y 5, en este caso es moda compuesta.

En muchos aspectos la moda se hace notar, aunque nos siempre es de manera numérica. Un ejemplo muy simple puede ser cuando un profesor analiza los resultados obtenidos de cierto grupo, por simple observación se puede ver cuál es la nota en la que más alumnos coincidieron.

Las medidas de variabilidad son "las que entregan información sobre la variación de la variable. Pretenden resumir en un solo valor la dispersión que tiene un conjunto de datos", Los estadísticos de variabilidad son los que nos permiten tener una idea de la dispersión de los datos respecto a algún valor promedio. Se refiere a la extensión de los datos de una distribución. Las medidas de dispersión más utilizadas son: Rango de variación, Varianza, Desviación estándar.

El rango es la diferencia entre el mayor valor de la variable y el menor valor de la variable.

Ejemplo:

29	31	35	39	39	40	43	44	44	52

Rango = 52 - 29 = 23

La varianza por otro lado es un proceso que se rige por una fórmula. Al obtener el valor de la varianza, por medio de una raíz cuadrada, obtenemos la deviación estándar. En los problemas estadísticos esta medida siempre va a requerir de la media aritmética, el valor de la desviación estándar siempre va a depender de la varianza.

$$\overline{X} = \frac{490+500+510+515+520}{5} = \frac{2535}{5} = 507$$

La varianza sería:

$$S^2 = \frac{(490-507)^2+(500-507)^2+(510-507)^2+(515-507)^2+(520-507)^2}{(5-1)}$$

$$S^2 = \frac{(-17)^2+(-7)^2+(3)^2+(8)^2+(13)^2}{4} = \frac{289+49+9+64+169}{4} = \frac{580}{4} = 145$$

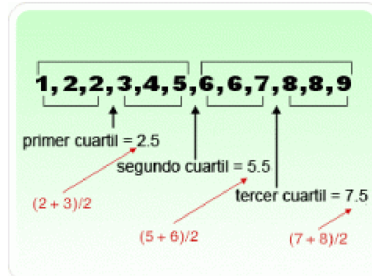
Por lo tanto la desviación estándar sería:

$$S = \sqrt{145} = 12.04 \cong 12$$

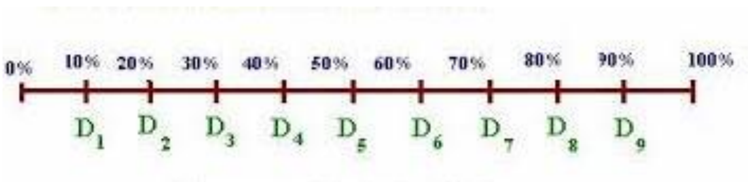
Ahora, dentro de los cuantiles tenemos tres: cuartiles, centiles y percentiles.

Los cuantiles "son aquellos valores de la variable que, ordenados de menor a mayor, dividen a la distribución en partes, de tal manera que cada una de ellas contiene el mismo número de frecuencias", es decir, son cantidades que dividen la distribución de datos en partes iguales, según sea el caso.

Los cuartiles son los valores que dividen en porcentajes 25% la totalidad de datos, Son 3 cuartiles. El segundo cuartil debe ser equivalente a la mediana y se representan con la letra Q. Los datos deben ordenarse estrictamente en forma creciente.



Deciles: Son los que dividen en forma equivalente la totalidad de datos, van de 10% y son nueve. El quinto cuartil equivale a la mediana. Se representan con la letra D.



Percentiles: Son 99, indican el valor de cada porcentaje, ya sea 1%, 2%, … Se representan con la letra P.

Para cada una de las tres, tenemos la fórmula para hallar su posición, aparte de eso se le aplica una segunda fórmula para encontrar el valor de cuantil, según se requiera.

En el caso de las asimetrías, como la de Yule Bowley o Karl pearson, los cuantiles son indispensables para la resolución y determinación de la asimetría. Se debe tener especial cuidado en las cantidades que ocupamos para que así las operaciones sean exactas y evitemos los errores. Se corre riesgo que alguna operación nos falle especialmente cuando en los problemas se nos pide encontrar distintas asimetrías o todas las medidas de dispersión y de variabilidad. Se debe tener especial cuidado en cada decimal y en el óptimo empleo de las formulas.

Existen muchas modalidades o circunstancias por las que estas importantes medidas pueden encontrarse. Ya sea con datos simples o agrupados, en tablas de frecuencia (con o sin intervalos) o solo en listado de datos. Cada una tiene su función y su fórmula. Para los datos agrupados en tablas de frecuencia a través de intervalos todo se torna más complejo. Cada una de las medidas toma una formula distinta y se pone mucha atención en todas las columnas y filas que esta posee, que por lo general son 9 u 10. Los procesos cambian.

Las formulas son las siguientes:

Media aritmética	**Mediana**	**Moda**
$\bar{x} = \frac{\sum_{i-1}^{n} x_i \cdot f_i}{n}$	$Me = L_i + \frac{\frac{N}{2} - F_{i-1}}{f_i} \cdot t_i$	$Mo = L_i + \frac{f_i - f_{i-1}}{(f_i - f_{i-1}) + (f_i - f_{i+1})} \cdot t_i$

Todos los símbolos representan una cantidad específica que debemos colocar para que la formula cobre vida y podamos obtener la cantidad requerida y exacta. El orden de las filas y las columnas es vital.

Hablemos de los coeficientes (que también son medidas). En estadística, cuando se desea hacer referencia a la relación entre el tamaño de la media y la variabilidad de la variable, se utiliza el coeficiente de variación. Este no es tan empleado, pero no deja de ser importante.

Los coeficientes de asimetría van a determinar si la asimetría de la distribución de los datos es positiva (a la derecha) o negativa (a la izquierda). Abordemos el coeficiente de Karl Pearson o solo coeficiente de Pearson. Esta depende en gran manera de las medidas de tendencia central pues básicamente se determina la relación que existe entre ellas. Para cada caso se rige por medio de una formula distinta la cual nos permite ser objetivos y reducir errores. Es una herramienta para los problemas estadísticos que requieren de resultados exactos y de análisis concretos.

Quizá las medidas mencionadas no son del todo comunes, pero si podemos hacer uso de ellas en el ámbito estadístico. De alguna manera, en algún punto de nuestra vida podemos hacer uso de ellas. Nos daremos cuenta que son un poco tediosas, pero nos resultaran de mucha utilidad.

Como seres humanos sabemos que la estadística, como ciencia, es de vital importancia y lo ha sido a lo largo del tiempo para el desarrollo social, cultural … Los científicos han realizado descubrimientos importantes en todos los ámbitos, ellos hacen uso de todas estas medidas y podemos decir que ellas son un granito de arena para el mundo científico y matemático.

Quizá las definiciones no son tan completas, pero tenemos idea de lo especial que son, Gracias a ellas podemos interpretar y obtener resultados que no tendríamos de no ser por ellas. Cada una de ellas merece especial cuidado. Todas son importantes.

Fuentes consultadas:

https://www.medwave.cl/link.cgi/Medwave/Series/MBE04/4934

https://rua.ua.es/dspace/bitstream/10045/19380/3/Tema%203-Estad%C3%ADstica.pdf

https://plataformaeducativauds.com.mx/assets/docs/files/asignatura/14f15eab271e2843e388ce4f42eced0c.pdf

II. Proporcionar ejemplos.

Proporcionar 2 ejemplos de medidas de tendencia central para datos agrupados.

Clase	Li	Ls	Fi	mc	Fi(mc)	fa	Contenido de:	Contenido a:
1	30	39	758	(30+39)/2=34.5	(758)(34.5)= 26151	758	1	758
2	40	49	635	(40+49)/2=44.5	(635)(44.5)=28257.5	1393	759	1393
3	50	59	104	(50+59)/2=54.5	(104)(54.5)=5668	1497	1394	1497
4	60	69	796	(60+69)/2=64.5	(796)(64.5)=51342	2293	1498	2293
5	70	79	435	(70+79)/2=74.5	(435)(74.5)=32407.5	2728	2294	2728
			∑=2728		∑=143826			

Ancho de clase:

A=Li (clase)-Li (clase n-1)

A= 40-30= 10

Media aritmética

$$:\bar{X} = \frac{1(mifi)}{n}$$

$$\bar{X} = \frac{143826}{2728}$$

$\bar{x}$= 52.72

Mediana:

Primero debemos sacar la posición de la mediana.

$$P.M. = \frac{n+1}{2}$$

$$P.M. = \frac{2728+1}{2}$$

P.M.= 1364.5

$$Med = Licm + \frac{n/2 - \sum ACM}{f(mediana)} * A$$

$$Med = 40 + \frac{\frac{2728}{2} - 758}{635} * 10$$

Med=49.5433

Moda=

$$Mod = LicmA \frac{ficmA - fiAcmA}{2ficmA - fiAcmA - fipcmA} * A$$

$$Mod = 60 \frac{796 - 104}{2(796) - 104 - 435} * 10$$

Mod= 66.571

Clase	Li	Ls	Fi	mc	Fi(mc)	fa	Contenido de:	Contenido a:
1	12	18	18	(12+18)/2=15	(18)(15)=270	18	1	18
2	20	26	12	(20+26)/2=23	(12)(23)=276	30	19	30
3	28	34	19	(28+34)/2=31	(19)(31)=589	49	31	49
4	36	42	29	(36+42)/2=39	(29)(39)=1131	78	50	78
5	44	50	10	(44+50)/2=47	(10)(47)=470	88	79	88
			∑=88		∑=2736			

Ancho de clase=

A= Li (clase)-Li (clase n-1)

A= 20-12= 8

Media aritmética=

$$:\bar{X} = \frac{1(mifi)}{n}$$

$$\bar{X} = \frac{2736}{88}$$

$$\bar{X} = 31.09$$

Mediana=

Posición de la mediana=

$$P.M. = \frac{n+1}{2}$$

$$P.M. = \frac{88+1}{2}$$

P.M.= 44.5

$$Med = Licm + \frac{n/2 - \sum ACM}{f(mediana)} * A$$

$$Med = 28 + \frac{88/2 - 30}{19} * 8$$

Med=33.8944

Moda=

$$Mod = LicmA \frac{ficmA - fiAcmA}{2ficmA - fiAcmA - fipcmA} * A$$

$$Mod = 36 \frac{29 - 19}{2(29) - 19 - 10} * 8$$

Mod=38.7584

Proporcionar 2 ejemplos de medidas de variabilidad para datos agrupados.

Xi	fi	Xi(fi)	Fi(xi-x)²
3	4	12	$(4)(3-7.86)^2=94.4784$
6	2	12	$(2)(6-7.86)^2=6.9192$
10	5	50	$(5)(10-7.86)^2=22.898$
12	3	36	$(3)(12-7.86)^2= 51.4188$
∑=	14	110	

Media aritmética=

$$:\bar{X} = \frac{1(Xifi)}{n}$$

$$=\frac{110}{14} = 7.86$$

Rango=

R= Xmáx-Xmin   (Los valores se toman de Xi)

=12-3 = 9

Varianza=

$$S^2 = \frac{1fi(xi-x)^2}{n}$$

$$S^2 = \frac{175.71}{14}$$ $S^2=12.55$

Desviación estándar=

$S=\sqrt{S^2} = 3.54$

Xi	fi	Xi(fi)	Fi(xi-x)²
5	19	95	$(19)(5-10.6285)^2=601.9202$
8	12	96	$(12)(8-10.6285)^2=82.9081$
11	16	176	$(16)(11-10.6285)^2=2.2081$
13	10	130	$(10)(13-10.6285)^2= 56.2401$
19	13	247	$(13)(19-10.6285)^2= 911.0661$
∑=	70	744	

Media aritmética=

$:\bar{X} = \frac{1(Xifi)}{n}$

$=\frac{744}{70}$ = 10.6285

Rango=

R= Xmáx-Xmin (Los valores se toman de Xi)

=19-5 =14

Varianza=

$$S^2 = \frac{1fi(xi-x)^2}{n}$$

$$S^2 = \frac{1654.3426}{70} = 23.6334$$

S=$\sqrt{S^2}$ =4.8614

III. Resuelve los siguientes ejercicios.

1.-Calcula la media aritmética, mediana, moda para los siguientes datos no agrupados: 6, 7, 8, 9, 10, 7, 8, 9, 7

Media aritmética=

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

$$=\frac{71}{9} = 7.89$$

Mediana=

Si queremos conocer la posición de la mediana debemos utilizar esta fórmula.

$$P.M. = \frac{n+1}{2}$$

P.M. = 9+1= 10/2= 5

(Ordenar las cantidades de mayor a menor)

6, 7, 7, 7, 8, 8, 9, 9, 10.

La mediana es 8 pues es el valor central. Comprobamos que el valor de la mediana está en la posición 5.

Moda=

La moda solo requiere observación.

6, 7, 7, 7, 8, 8, 9, 9, 10.

Como vemos que la cantidad que más veces se repite es 7, esa es nuestra moda.

2.- Calcula el rango, varianza y desviación estándar para los siguientes datos: 46, 55, 50, 47, 52.

$$:\overline{X} = \frac{46+55+50+47+52}{5} =50$$

Rango=

R= Xmáx-Xmin

=55-46 =9

Varianza=

$$S^2 = \frac{(46-50)^2 + (55-50)^2 + (50-50)^2 + (47-50)^2 + (52-50)^2}{5-1}$$

$$S^2 = \frac{(-4)^2 + (5)^2 + (0)^2 + (-3)^2 + (52-50)^2}{5-1}$$

$$S^2 = \frac{54}{4}$$

S^2=13.5

Desviación estándar=

S=$\sqrt{S^2}$=$\sqrt{13.5}$

S=3.6742

3.- La siguiente tabla representa la calificación de matemáticas I de un grupo de alumnos de la secundaria:

Calificación	Número de alumnos
4	3
5	6
6	9
7	16
8	12
9	5
10	3
	∑=54

a) Calcula la calificación promedio

$$:\bar{X} = \frac{4+4+4+5+5+5+5+5+5+6+6+6+6+6+6+6+6+6+7+7+7+7+7+7+7+7+7+7+7+7+7+7+7+7+8+8+8+8+8+8+8+8+8+8+8+8+9+9+9+9+9+10+10+10}{54}$$

$=\frac{379}{54}$ =7.0185

b) Calcula el porcentaje de alumnos que aprobaron la materia.

De los 54 alumnos que hay solo 45 aprobaron la materia.

¿Cuál es el porcentaje de aprobados?

Resolvemos por medio de una regla de tres simple.

54 ———————— 100%

45 ———————— X

$$X = \frac{(45)(100)}{54}$$

$$X = \frac{(4500)}{54}$$

X= 83.3%