Aceves Jenaro Vidal Gladiolo
6IM7

11|II|2022 Sección A
Sesión #5

Demostración $\bar{x} \geq G \geq H$ $[\bar{x} \geq G$ inducción$]$

Proponemos dos números a y b. Las relaciones entre estos dos números pueden ser las siguientes: $a=b$, $a>b$, $b>a$ (ley de tricotomía).

Para fines de este curso, definimos que la diferencia debe de ser mayor o igual a 0, $a-b \geq 0$. (tomaremos $a \geq b$)

$$\therefore \quad (a-b)^2 \geq 0 \Rightarrow a^2 - 2ab + b^2 \geq 0$$

dejamos cuadráticos de un lado de la igualdad (los aislamos), y completamos un trinomio cuadrado perfecto, sumando a ambos lado 2ab.

$$a^2 - 2ab + b^2 \geq 0 \quad \Big| \quad a^2 + 2ab + b^2 \geq 2ab + 2ab$$
$$a^2 + b^2 \geq 2ab \quad \Big| \quad (a+b)^2 \geq 4ab$$

convertimos nuestra igualdad cuadrática a una lineal, elevando ambos lados a la potencia $1/2$

$$\left[(a+b)^2\right]^{1/2} \geq \left[4ab\right]^{1/2} \quad \Big| \quad a+b \geq 2\sqrt{ab}$$

aislamos el término $\sqrt{ab}$, y notamos que los integrantes son respectivamente la media aritmética y la geométrica.

$$\frac{a+b}{2} \geq \sqrt{ab} \Rightarrow \bar{x} \geq G \quad \text{Q.E.D.}$$

$\therefore \bar{x} \geq G$ [para dos números]

A través de la trigonometría y geometría analítica, podemos demostrar que para los dos números a, b, la media armónica es la distancia dentre el foco y la directriz de una parábola que contiene a ambos números.

ej Los salarios reportados ante el SAT de 4 personas que trabajan en una pequeña empresa fueron los siguientes: \$5'000, \$6'000, \$6'500 y \$30'000.
Calcule las 5 MTC.

$$\bar{x} = \frac{(5+6+6.5+30)\,1000}{4} = \underline{11'875}$$

$$G = \sqrt[4]{(5)(6)(6.5)(30)(1000)^4} = 1000\sqrt[4]{30^2 \cdot 6.5} = \underline{8'745.5870}$$

$$H = 4\left[\left(\tfrac{1}{5}+\tfrac{1}{6}+\tfrac{2}{13}+\tfrac{1}{30}\right)\tfrac{1}{1000}\right]^{-1} = 4'000\left[36/65\right]^{-1} = 4'000\,(65/36) = \underline{7'222.2222}$$

M_o : no hay.

M_d : $M_{d\,pos} = \frac{4+1}{2} = 2.5 \lessgtr \begin{matrix} 2^a \\ 3^a \end{matrix}$

2^a pos : 6'000 $\quad$ 3^a pos : 6'500

$$M_d = \frac{6'000 + 6'500}{2} = \underline{6'250}$$

Aceus Jonavo Vieles Gbalieb 6IV17

11|II|2022 Sección A
Sesión #5

Para que una muestra sea buena, debe haber sido seleccionada al azar y tener características apropiadas.
Un grave problema que presenta la media aritmética es que es muy sensible a los valores extremos de un conjunto de datos; cuando eso sucede es más aconsejable emplear a la mediana.

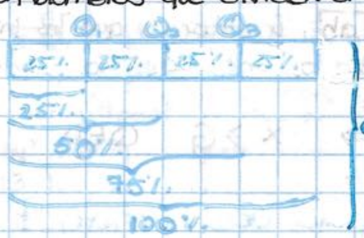
La media armónica trata de equilibrar los grandes "baches" de los datos, un ejemplo de esto es en el ejercicio anterior (salarios SAT) con los datos \$5'000 y \$30'000.

Medidas de posición.

Sirve para describir la localización específica de un dato en relación con el resto de la muestra o población. Las medidas de posición o de ubicación más populares son las siguientes:

i. Cuartiles : Q_K

Son números que dividen a los datos ordenados en 4 partes.

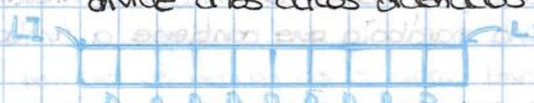


LI : límite inferior
LS : límite superior

$Q_K \Rightarrow K = 1, 2, 3.$

ii. Deciles : D_K

divide a los datos ordenados en 10 partes.



$D_K \Rightarrow D = 1, 2, 3, 4, 5, 6, 7, 8, 9.$

iii. Percentiles : P_K

divide a los datos ordenados en 100 partes.

$P_K \Rightarrow P = 1, 2, 3, \ldots, 97, 98, 99.$

Arcos Ferrero Víctor Ghalib — 11 | II | 2022 — Sección A
6IM7 — Sesión #5

La mediana también puede ser utilizada como medida de posición, siendo Q_2, D_5 o P_{50}.

Todas las posiciones son relativas al límite inferior y nos sirven de referencia para la frecuencia acumulada (Fa) o viceversa, por ejemplo:

$$M_d = D_5 = Q_2 = P_{50} \Leftrightarrow Q_1 = P_{25},\ Q_3 = P_{75},\ D_4 = P_{40}$$

→ **Procedimiento Batiziano para obtener las medidas:**

$Md_{pos} = \frac{n+1}{2}$

1. Ordenar los datos.
2. Si se trata de: Q_2, D_5 o P_{50}, igualamos Md.
3. Encontrar la posición del cuantil.*

$$Q_{K_{pos}} = \frac{nK}{4} \qquad D_{K_{pos}} = \frac{nK}{10} \qquad P_{K_{pos}} = \frac{nK}{100}$$

4. Si la posición encontrada no es un entero, se redondea al entero más próximo.
5. Si la fracción es 0.05, se obtiene la media entre el valor anterior y posterior.

* Nombre genérico.

ej. El siguiente conjunto de datos ordenados representa una muestra de los rendimientos de gasolina, en Km/l, reportados por una compañía que se dedica a la compraventa de autos usados.

6.5	7.2	7.5	8.2	8.7	9.1	9.3	9.5	10.1	11.4	11.8	11.9	12.0	13.1	13.5	14.2	14.5	14.7
1°	2°	3°	4°	5°	6°	7°	8°	9°	10°	11°	12°	13°	14°	15°	16°	17°	18°

Encontrar las siguientes medidas de posición:

1. P_{10}: $P_{10_{pos}} = \frac{18(10)}{100} = 1.8 \approx 2 \Rightarrow 2^a_{pos} = 7.2 \Rightarrow P_{10} = 7.2\ Km/l$

2. Q_1: $Q_{1_{pos}} = \frac{18(1)}{4} = 4.5 \begin{smallmatrix} \nearrow 5^a \\ \searrow 4^a \end{smallmatrix} \Rightarrow \begin{matrix} 5a_{pos} = 8.7 \\ 4a_{pos} = 8.2 \end{matrix} \Rightarrow Q_1 = \frac{8.7+8.2}{2} = 8.45\ Km/l$

3. Q_2: $Md_{pos} = \frac{n+1}{2} = \frac{19}{2} = 9.5 \begin{smallmatrix} \nearrow 10^a \\ \searrow 9^a \end{smallmatrix} \Rightarrow \begin{matrix} 10a_{pos} = 11.4 \\ 9a_{pos} = 10.1 \end{matrix} \Rightarrow Q_2 = \frac{11.4+10.1}{2} = 10.75\ Km/l$

4. Q_3: $Q_{3_{pos}} = \frac{18(3)}{4} = 13.5 \begin{smallmatrix} \nearrow 14^a \\ \searrow 13^a \end{smallmatrix} \Rightarrow \begin{matrix} 14a_{pos} = 13.1 \\ 13a_{pos} = 12.0 \end{matrix} \Rightarrow Q_3 = \frac{13.1+12.0}{2} = 12.55\ Km/l$

5. P_{90}: $P_{90_{pos}} = \frac{18(90)}{100} = 16.2 \to 16^a \Rightarrow 16a_{pos} = 14.2 \Rightarrow P_{90} = 14.2\ Km/l$

Aceves Serrano Victor Ghalieb 11|II|2022 Sección A
6IM7 Sesión #5

Resumen con 5 números.

(descripción -muy- eficaz de un conjunto de datos).

a) LI	6.5	c) Q_2	10.75	e) LS	14.2
b) Q_1	8.45	d) Q_3	12.55		

Un diagrama de caja representa los datos, según la manera en la que se reparten los datos. En las cajitas está concentrado el 50% de los datos centrales, y por lo tanto, es estadística.

diagrama de caja con brazos ... o con bigote.

Indicadores que utilizan a los cuartiles:

- Rango Intercuartil: RI
  $Q_3 - Q_1$ $\quad 12.55 - 8.45 = 4.1$
- Rango Semi-intercuartil: RSI
  $\frac{1}{2}(Q_3 - Q_1) = \frac{1}{2} RI$ $\quad \frac{1}{2}(4.1) = 2.05$
- Cuartil Intermedio: QI
  $\frac{1}{2}(Q_1 + Q_3)$ $\quad \frac{1}{2}(12.55 + 8.45) = 10.5$

* también es una medida de tend. central.

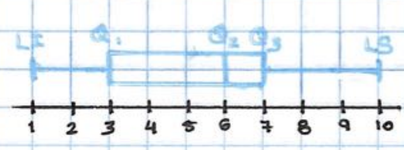
ej. Para el siguiente conjunto de valores, encontrar los 3 cuartiles.

x	f	F
1	4	4
2	3	7
3	3	10
4	4	14
5	5	19
6	7	26
7	5	31
8	5	36
9	3	39
10	1	40

a) Q_1: $Q_{1_{pos}} = \frac{40(1)}{4} = 10^a \rightarrow 10^a_{pos} = 3 \Rightarrow Q_1 = 3$

b) Q_2: M_d: $M_{d_{pos}} = \frac{40+1}{2} = 20.5 \lessgtr {21° \atop 20°} \Rightarrow {21°_{pos} = 6 \atop 20°_{pos} = 6} \Rightarrow Q_2 = 6$

c) Q_3: $Q_{3_{pos}} = \frac{40(3)}{4} = 30 \rightarrow 30^a_{pos} = \Rightarrow Q_3 = 7$



RI = 7 - 3 = 4

RSI = $\frac{1}{2}(4) = 2$

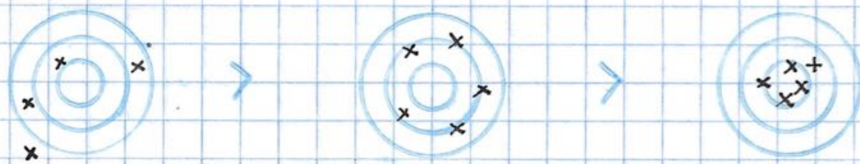
$Q_I = \frac{1}{2}(7+3) = 5$

Norma

Aceves Serrano Víctor Ghalieb 11|II|2022 Sección A
6IM7 Sesión #5

Medidas de dispersión

Estas medidas indican el grado de dispersión o variabilidad del conjunto de datos. Estas medidas serán mayores cuando los datos estén más desagregados y serán menores cuando estén más cercanamente agrupados.



Los medidas de dispersión a usar en el curso, son:

i. **Amplitud o rango:** $R = LS - LI$

ii. **Varianza muestral:** $S^2 = \left(\frac{1}{n-1}\right)\left(\Sigma\left[(x-\bar{x})^2\right]\right) \quad \forall n \in n \leq 30$

ii'. **Varianza poblacional:** $\sigma^2 = \left(\frac{1}{n}\right)\left(\Sigma\left[(x-\bar{x})^2\right]\right) \quad \forall n \in n \geq 31$

iii. **Desviación estandar muestral:** $S = \sqrt{\left(\frac{1}{n-1}\right)\Sigma\left[(x-\bar{x})^2\right]} \quad \forall n \in n \leq 30$

iii'. **Desviación estandar poblacional:** $\sigma = \sqrt{\left(\frac{1}{n}\right)\Sigma\left[(x-\bar{x})^2\right]} \quad \forall n \in n \geq 31$

ej. Dos estudiantes reportan las siguientes calificaciones en su 1er ex. de licenciatura:

x	$(x-\bar{x})$	$(x-\bar{x})^2$
5	-2	4
5	-2	4
6	-1	1
6	-1	1
7	0	0
8	1	1
9	2	4
10	3	9
$\Sigma x = 56$	$\Sigma(x-\bar{x}) = 0$ *	$\Sigma[(x-\bar{x})^2] = 24$

x	$(x-\bar{x})$	$(x-\bar{x})^2$
3	-4	16
6	-1	1
6	-1	1
7	0	0
7	0	0
8	1	1
9	2	4
10	3	9
$\Sigma x = 56$	$\Sigma(x-\bar{x}) = 0$	$\Sigma[(x-\bar{x})^2] =$

$S_1 < S_2$

$\bar{x}_1 = \frac{56}{8} = 7$

$s_1^2 = \frac{24}{8-1} = 3.4286$

$S = 1.8516 \qquad R_1 = 10-5 = 5$

$\bar{x}_2 = \frac{56}{8} = 7$

$s_2^2 = \frac{32}{7} = 4,5714$

$S_2 = 2.1381 \qquad R_2 = 10-3 = 7$

* En una tabla bien realizada, siempre $\Sigma(x-\bar{x}) = 0$.