apartir de aquí

Mercado Vargas Melisa 6IM7
Sección B 11/02/22

¿Porqué $\bar{x} \geq G \geq H$?

Sea a, b; $a=b$, $a>b$ ó $a<b$ (ley de tricotomía)
$a-b \geq 0$, para garantizar que esa diferencia sea ≥ 0. entonces $(a-b)^2$. $a^2-2ab+b^2 \geq 0$

$a^2+b^2 \geq 2ab$

$a^2+2ab+b^2 \geq 2ab+2ab$

$((a+b)^2 \geq 4ab)^{1/2}$

$a+b \geq 2\sqrt{ab}$

$\frac{a+b}{2} \geq \sqrt{ab}$

$\bar{x} \geq G$ (para dos números)

Ejercicio. Dos salarios reportados ante el SAT de 4 personas que trabajan en una pequeña empresa fueron los siguientes: \$5000 ; \$6,000 ; \$6,500 ; \$30,000. Calcule las 3 medidas de tendencia central.

$\bar{x} = \frac{5,000+6,000+6500+30000}{4} = \$11\,875$

$G = \sqrt[4]{(5000)(6000)(6500)(30000)} = \$8745.38$

$H = \frac{4}{\frac{1}{5000}+\frac{1}{6000}+\frac{1}{6500}+\frac{1}{30,000}} = \$7222.22$

Mediana $Md = \frac{4+1}{2} = 2.5$ entre 2 y 3

$M = \frac{6000+6500}{2} = \$6\,250$

Moda: No hay $\emptyset$. Es Amodal

Mercado Vargas Melsa 6IM7
Sección B 11/02/22

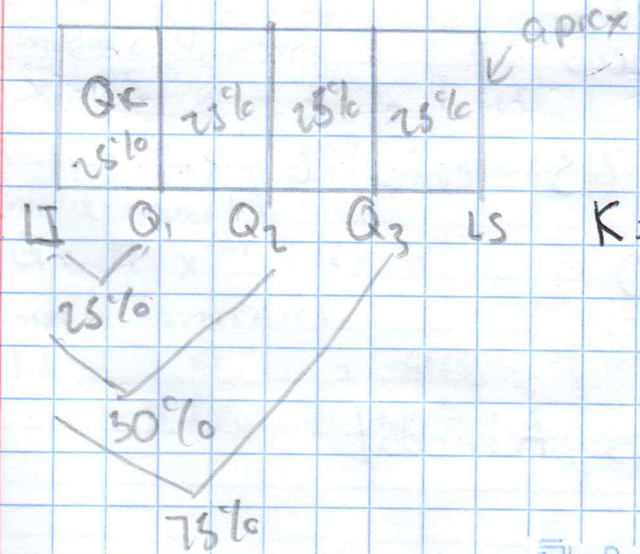
Para que una muestra sea buena, debió haber sido tomada al azar y tener características apropiadas. ~~Para no de~~

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 harmónica trata de equilibrar los grandes "bachos" de los datos ($5000 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 sigs.

i) Cuartiles (Q_K) (De cuartos): Son números que dividen a los datos ordenados en 4 partes.

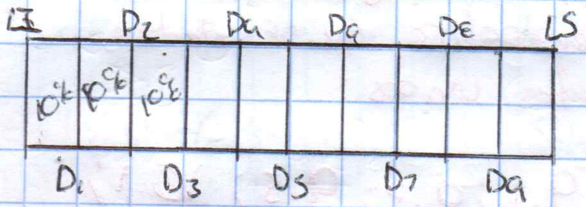


LI: Límite inferior
LS: Límite superior

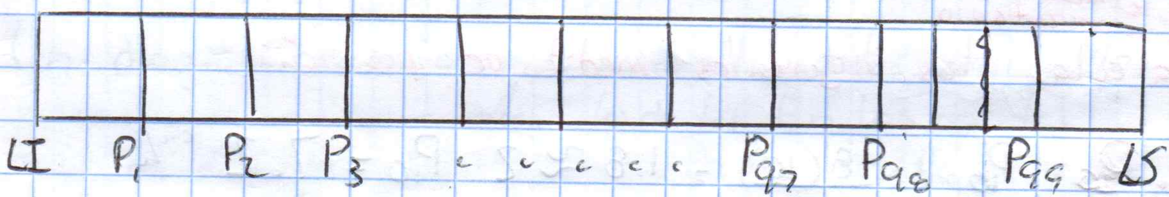
K: 1, 2, 3

Mercado Vargas Melsa 6IM7
Sección B 11/02/22

2) Deciles (D_{10}) $K = 1, 2, 3, 4, 5, 6, 7, 8, 9$



3) Percentiles (Percentiles) P_K $K = 1, 2, 3, \dots, 99$



La mediana tambien puede ser tambien una medida de posición, siendo Q_2, D_5 o P_{50}.

$Md = D_5 = Q_2 = P_{50}$

$Q_1 = P_{25}$ $D_4 = P_{40}$

$Q_3 = P_{75}$

Todo es relativo al LI, y con base a él, de ahí para arriba. Y sirve con fa (P)

Procedimiento "Batiz" para obtener las medidas

1) Ordenar los datos $Md_{pos} = \frac{n+1}{2}$

2) Si se trata de: Q_2, D_5, P_{50}, se saca Md.

3) Encontrar la posición del "Cuantil" (nombre genérico)

$Q_{K\,pos} = \frac{nK}{4}$ $D_{K\,pos} = \frac{nK}{10}$ $P_{K\,pos} = \frac{nK}{100}$

4) Si la posición encontrada no es un entero, entonces la posición se redondea al entero más próximo.

5) Si la fracción es 0.5, se toman los valores anterior y posterior a la posición y se calcula la media de estos dos

Mercado Vargas Melisa 6IM7
Sección B 11/02/22

Ejemplo. El sig. conjunto de datos ordenados representa una muestra de los rendimientos de gasolina en Km/L reportados por una compañía que se dedica a la compra y venta de autos usados:

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

no se calculan, se encuentran :o

Calcular los siguientes medidas de posición:

1. P_{10} $\quad P_{10\,pos} = \frac{18(10)}{100} = 1.8 \approx 2 \quad P_{10} = 7.2$ km/L

4° y 5°

2) Q_1 $\quad Q_{1\,pos} = \frac{18(1)}{4} = 4.5 \quad Q_1 = \frac{8.2 + 8.7}{2} = 8.45$ km/L

9; 10°

3.- Q_2 $\quad \cancel{Md_{pos} = \frac{n}{2} = 18} \quad Md_{pos} = \frac{n+1}{2} = 9.5 \quad M = \frac{10.1 + 11.4}{2} = 10.75$

$Q_2 = 10.75$ km/L

13° y 14°

4= Q_3 $\quad Q_{3\,pos} = \frac{18(3)}{4} = 13.5 \quad Q_3 = \frac{12.0 + 13.1}{2} = 12.55$ km/L

16

5.) P_{90} $\quad P_{90\,pos} = \frac{18(90)}{100} = 16.2 \quad P_{90} = 14.2$ km/L

Mercado Vargas Melsa 61147
Sección B 11/02/22

Resumen con 5 números

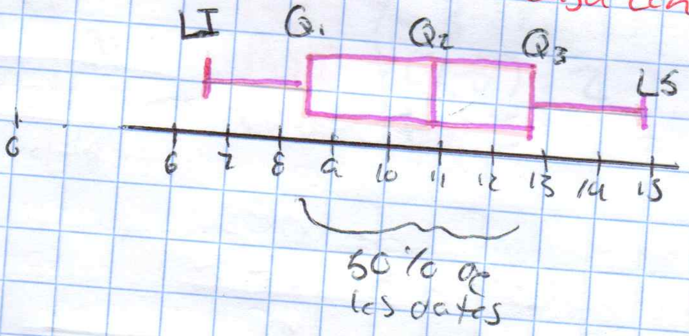
Resulta ser un medio muy eficaz para la descripción de un conjunto de datos y son:

a) LI: 6.3
b) Q_1 8.45
c) Q_2 10.75
d) Q_3 12.55
e) LS 14.7

Son datos continuos

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 esa longitud es en ~~estado~~ estadístico

Diagrama de caja (o caja con brazos / bigotes)



Mercado Vargas Melissa
6IM7 11/02/22

Indicadores que utilizan a los cuartiles

1) Rango Intercuartil RI $Q_3 - Q_1$ $12.55 - 8.45 = 4.1$

2) Rango Semiintercuartil RSI $\frac{1}{2}(Q_3 - Q_1) = 2.05$

3) Cuartil Intermedio QI $\frac{Q_1 + Q_3}{2} = 10.5$

↗

Tambien es una MTC

Otras medidas:

Media Ponderada

Media Movil

Mercado Vargas Melse
6IM7 Seccion B 11/02/22

Ejercicio. Para el sig. conjunto de datos frecuenciales, encuentre los 3 cuartiles y elabore su diagrama de caja.

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

a) Q_1 $\quad Q_1\,pos = \frac{40(1)}{4} = 10\ pos$

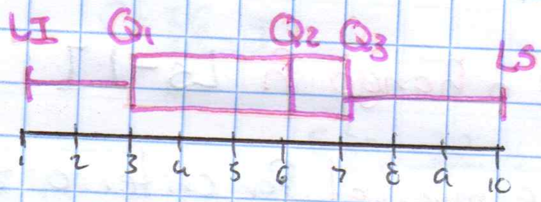
$Q_1 = 3$

b) $Q_2 = Md$ $\quad Md\,pos = \frac{40+1}{2} = 20.5$ $\rightarrow 20$ $\rightarrow 21$

$Q_2 = Md = \frac{6+6}{2} = 6$

c) Q_3 $\quad Q_3\,pos = \frac{40(3)}{4} = 30°$ $\quad Q_3 = 7$

d)



$RI = 7 - 3 = 4$

$RSI\ \frac{1}{2}(7-3) = 2$

$QI\ \frac{2+3}{2} = 3$

Mercado Vargas Melss
6IM7 Sección B 11/02/22

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 disgregados y serán menores cuando estén más cercanamente agrupados.

Tiro en blanco

Ver qué tan lejos están según la media

Las medidas de dispersión que emplearemos en este curso serán

1) Amplitud o Rango: R $Ls - LI$

2) Varianza muestral Por criterio si $n \leq 30$ se le dice muestral y se representa como s^2 $s^2 = \frac{\sum(x-\bar{x})^2}{n-1}$

3) Desviación estándar muestral $n \leq 30$ $s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$

Para una población se considera que $n > 30$

4) Varianza Poblacional $\overset{\text{sigma}}{\sigma^2} = \frac{\sum(x-\bar{x})^2}{n}$

3) Desviación estándar poblacional $\sigma = \sqrt{\frac{\sum(x-\bar{x})^2}{n}}$

La varianza siempre es ≥ 0, si es un conjunto de datos parejo sí puede ser 0

Mercado Vargas Melisa 6IM7
Sección B 11/02/22

Ejemplo Dos estudiantes reportan el siguiente conjunto de calificaciones en su primer semestre de Licenciatura.

Estudiante 1

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
$\Sigma x = 56$	$\Sigma = 0$	$\Sigma = 24$

Estudiante 2

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
$\Sigma x = 56$	$\Sigma\ 0$	32

$$\bar{x} = \frac{\Sigma x}{n} = 7$$

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

$$s_1^2 = \frac{\Sigma(x-\bar{x})^2}{n-1} = \frac{24}{8-1} = 3.43$$

$$s_2^2 = \frac{32}{8-1} = 4.5714$$

$$s_1 = \sqrt{3.43} = 1.8520$$

$$s_2 = \sqrt{4.5714} = 2.1380$$

mientras más grande es, mayor dispersión

$R_1 = 10-5 = 5$

$R_2 = 10-3 = 7$

$s_1 > s_2$