Arreves Serrano Victor Ghalieb 14|II|2022 Sección A

Sesión #6

- Recordatorio:

Estadística descriptiva
- MTC — $\bar{x}, G, H, Md, Mo$
- M. Posición — Q_K, D_K, P_K
- M. Dispersión — $R, s^2, \sigma^2, s, \sigma$

Cuando tenemos un conj. d/datos con 30 o menos elementos, hablamos de un conj. d/datos muestral, mientras que al tener una con 31 o más elementos, hablamos de una poblacional.
La diferencia entre uno conjunto de datos muestral (s) y poblacional (σ) radica en la manera en la que interpretamos el sesgo de la varianza.

La varianza muestral (s^2) es un estimador sin sesgo de la varianza poblacional (σ^2), esto significa que los valores de s^2 tienden a igualar el valor de σ^2, en lugar de tender, de manera sistemática, a sobreestimar o subestimar σ^2.

Es decir: cuando dividimos $\Sigma(x-\bar{x})^2$ entre n, la hacemos tendenciosa.

$$s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1} \qquad \sigma^2 = \frac{\Sigma(x-\bar{x})^2}{n}$$

ej. Utilizando las siguientes calificaciones, encontrar la varianza y desviación estandar.

x	$x-\bar{x}$	$(x-\bar{x})^2$
10	0.1	0.01
10	0.1	0.01
10	0.1	0.01
10	0.1	0.01
10	0.1	0.01
10	0.1	0.01
10	0.1	0.01
10	0.1	0.01
10	0.1	0.01
9	-0.9	0.81
$\Sigma x = 99$	$\Sigma(x-\bar{x}) = 0$	$\Sigma(x-\bar{x})^2 = 0.9$

$\bar{x} = \frac{99}{10} = 9.9$

$\Rightarrow n = 10 \Rightarrow$ Muestral ☒ Poblacional ☐

$\therefore s^2 = \frac{0.9}{(10-1)} = \frac{0.9}{9} = \frac{9}{9 \cdot 10} = \frac{1}{10} = 0.1$

$s^2 = 0.1$

$s = \sqrt{0.1} = \frac{\sqrt{10}}{10} = 0.3162$

$s = 0.3162$

Este método para calcular las varianzas puede tomarnos una cant. considerable de tiempo, por lo que podemos realizar un análisis de nuestras fórmulas y conseguir formulas compactas como el "método corto de la varianza".

Aceves Serrano Victor Ghalieb 14|II|2022 Sección A
61M7 Sesión #6

Método corto de la varianza

Conocemos nuestras expresiones para varianza:

$$s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1} \quad y \quad \sigma^2 = \frac{\Sigma(x-\bar{x})^2}{n},$$

pero tambien sabemos que: $\bar{x} = \frac{\Sigma(x)}{n}$. Notemos, de manera inicial, que dentro del numerador de las varianzas tenemos un binomio elevado al cuadrado. Podemos enfocarnos en este, y expandirlo.

$$\Sigma(x-\bar{x})^2 = \Sigma\left(x^2 - 2x\bar{x} + \bar{x}^2\right) \leftarrow \text{distribuimos } "\Sigma".$$

$$= \Sigma(x^2) - 2\Sigma(x\cdot\bar{x}) + \Sigma(\bar{x}^2) \leftarrow \text{sustituimos } \bar{x} = \frac{\Sigma(x)}{n}$$

$$= \Sigma(x^2) - 2\underbrace{\left[\Sigma(x)\cdot\frac{\Sigma(x)}{n}\right]}_{a\cdot a = a^2} + \underbrace{\Sigma\left[\left(\frac{\Sigma(x)}{n}\right)^2\right]}_{x\cdot a = x\cdot 1\cdot a;\ x\cdot\frac{a}{n} = x\cdot\frac{1}{n}\cdot a}$$

$$= \Sigma(x^2) - 2\frac{[\Sigma(x)]^2}{n} + \Sigma\left[\left(\frac{1}{n^2}\right)\right]\cdot[\Sigma(x)]^2$$

$$\sum_{i=1}^{n}(1) = n;\ \Sigma\left(\tfrac{1}{n^2}\right) = \Sigma\left[1\cdot\tfrac{1}{n^2}\right] = \Sigma(1)\cdot\Sigma\left(\tfrac{1}{n^2}\right)$$

$$= \Sigma(x^2) - 2\frac{[\Sigma(x)]^2}{n} + \frac{n}{n^2}(\Sigma x)^2$$

$$= \Sigma(x^2) - 2\frac{(\Sigma x)^2}{n} + \frac{(\Sigma x)^2}{n}$$

$$= \Sigma(x^2) - \frac{(\Sigma x)^2}{n} = \underbrace{\frac{1}{n}\left(n\Sigma(x^2) - (\Sigma x)^2\right)}_{\text{Numerador.}}$$

Una vez analizado el numerador, hemos conseguido expresarlo únicamente en términos de n, $\Sigma(x^2)$ y $\Sigma(x)$. Si insertamos esta nueva expresión dentro de nuestras fórmulas de varianza, tenemos:

$$s^2 = \frac{1}{n}\cdot\frac{n\Sigma(x^2) - (\Sigma x)^2}{n-1} \qquad \sigma^2 = \frac{1}{n}\cdot\frac{n\Sigma(x^2) - (\Sigma x)^2}{n}$$

Distribuimos $\frac{1}{n}$, separamos fracciones y obtenemos:

$$s^2 = \frac{n\Sigma(x^2)}{n(n-1)} - \frac{(\Sigma x)^2}{n(n-1)} \qquad \sigma^2 = \frac{n\Sigma(x^2)}{n^2} - \frac{(\Sigma x)^2}{n^2}$$

$$s^2 = \frac{\Sigma(x^2)}{n-1} - \frac{(\Sigma x)^2}{n^2-n} \qquad \sigma^2 = \frac{\Sigma(x^2)}{n} - \left(\frac{\Sigma x}{n}\right)^2$$

$$\sigma^2 = \frac{\Sigma(x^2)}{n} - \bar{x}^2$$

$\therefore$ **método corto:** $s^2 = \frac{\Sigma x^2}{n-1} - \frac{\Sigma^2 x}{n^2-n}$ $\quad \sigma^2 = \frac{\Sigma x^2}{n} - \bar{x}^2$

Arévalo Serrano Víctor Ghalieb
6IM7

De manera análoga, cuando hablamos de datos frecuenciales, tenemos:

$$s^2 = \frac{\Sigma(f\cdot x^2)}{n-1} - \frac{(\Sigma(f\cdot x))^2}{n^2-n} \qquad \sigma^2 = \frac{\Sigma(f\cdot x^2)}{n} - \bar{x}^2$$

ej.

x	f	$f\cdot x$	$f\cdot x^2$
1	4	4	$4(1)^2 = 4$
2	3	6	$3(2)^2 = 12$
3	3	9	$3(3)^2 = 27$
4	4	16	$4(4)^2 = 64$
5	5	25	$5(5)^2 = 125$
6	7	42	$7(6)^2 = 252$
7	5	35	$5(7)^2 = 245$
8	5	40	$5(8)^2 = 320$
9	3	27	$3(9)^2 = 243$
10	1	10	$1(10)^2 = 100$
Σ	40	214	1392

$40 > 31 \Rightarrow \sigma$

$\bar{x} = \frac{214}{40} = \frac{107}{20}$

$\sigma^2 = \frac{1392}{40} - \left(\frac{107}{20}\right)^2 = 6.1775$

$\sigma = \sqrt{\frac{2471}{400}} = 2.4855$

muy alta, existe mucha dispersión en los datos.

ej.

7	7	8	7	9
8	8	8	8	9
8	10	8	10	7
8	6	7	10	7
8	9	7	7	7

x		f	F	$\%Fa$	$f\cdot x$	$f\cdot x^2$
6	I	1	1	$\frac{100}{25}\cdot 1 = 4$	6	36
7	~~IIII~~ IIII	9	10	$4\cdot 10 = 40$	63	441
8	~~IIII~~ IIII	9	19	$4\cdot 19 = 76$	72	576
9	III	3	22	$4\cdot 22 = 88$	27	243
10	III	3	25	$4\cdot 25 = 108$	30	300
Σ					198	1596

Rango = LS − LI : 10 − 6 = 4

RI = $Q_3 - Q_1$: 8 − 7 = 1

RSI = $\frac{1}{2}$ RI : $\frac{1}{2}(1) = \frac{1}{2}$

$Q_x = \frac{1}{2}(Q_1 + Q_3) = \frac{1}{2}(8+7) = 7.5$

Q_1: $Q_{1pos} = \frac{25(1)}{4} = 6.25 \approx 6 \Rightarrow Q_1 = 7$

Md: $(25+1)/2 = 13 \Rightarrow Md = 8$ → posición

Q_3: $Q_{3pos} = \frac{25(3)}{4} = 18.75 \approx 19 \Rightarrow Q_3 = 8$

$s^2 = \frac{1596}{25-1} - \frac{198^2}{25^2 - 25} = \frac{29}{25} \approx 1.16$

$s = \sqrt{\frac{29}{25}} = \frac{\sqrt{29}}{5} \approx 1.0770$