

ESTADÍSTIKA ETA DATUEN ANALISIA

**V: Aldagai bakunaren deskribapena:
alborapena eta kurtosia**

Egilea: Josemari Sarasola



Gizapedia

gizapedia.hirusta.io

5.1 Itxura: alborapena eta kurtosia

5.2 Alborapena

5.2.1 Definizioa: datu kopurua batezbestekotik alde banatara

5.2.2 Definizioa: batezbestekoa eta mediana alderatuz

5.2.3 Alborapen neurriak

5.2.3.1 Fisher koefizientea

5.2.3.2 Bowley koefizientea

5.3 Kurtosia

5.3.1 Kurtosi-tipologia

5.3.2 Kurtosi neurriak

5.3.2.1 Pearson koefizientea

5.3.2.2 Moors koefizientea

5.3 Ariketak: alborapena eta kurtosia

1. Eguneko ekoiztutako pieza kopurua jaso da lantegi batean:

26-24-29-30-21-19-16-32-36-21-20-42-50-28-24

- Datuak puntu-diagrama batean jarri eta adierazi datuek alborapen nabarmena erakusten duten.
- Batezbestekoa eta mediana kalkulatu, adierazi norako alborapena dagoen.
- Batezbestekotik gora eta beheko datuak zenbatuz, adierazi norako alborapena dagoen.
- Fisher alborapen-koefizientea kalkulatu datu guztiak hartuta eta interpretatu.
- Fisher alborapen-koefizientea birkalkulatu bi datu handienak kenduta eta emaitza interpretatu.
- Estatistikan, simetria datu-multzoen propietate komenigarria da. Simetria lortzearen datuen aldakuntzarako funtzioak erabiltzen dira. Horietako bat erro karratu funtzioa izaten da. Datuei aldakuntza hori aplikatuz, alborapena nabarmen jaisten al da? Fisher koefizientea kalkulatu horretarako.

2. Hainbat familiaren errentari buruzko datuak jaso eta tartetan bildu dira:

Errentak	Familiak
200-300	126
300-400	459
400-500	896
500-600	367
600-700	292
700-800	85
	2225

- Batezbestekoa eta mediana kalkulatu eta horretan oinarrituz adierazi norako alborapena dagoen.
- Bowley alborapen-koefizientea eta Moors kurtosi-koefizientea kalkulatu eta interpretatu.
- Emaitza guztiak interpretatu banakuntza normalaren aplikagarritasunari buruz.

3. Autobus batek goizeko ibilbide bat egiteko behar dituen denborak jaso dira (minututan):

25-28-32-24-27-35-38-19-27-27-25-26-29-35-42

18-20-22-24-23-25-26-32-29-27-20-21-22-25-26-26

- Datuatarako grafiko egokia eratuz eta interpretatuz, azter ezazu datuen eredu moduan banakuntza normala egokia den.
- Neurri sendoak erabiliz, egiazta itzazu aurreko ondorioak.

4. Test batean ikasle batzuek lortutako kalifikazioak bildu dira:

Kalifikazioak	Ikasleak
20-30	8
30-40	26
40-50	86
50-60	32
60-70	14
	166

Datuetan jasotzen den informazio guztia kontuan hartzen duten neurriak kalkulatu, azter ezazu datu horietarako banakuntza normala egokia izan daitekeen.

5. 30 urteko pinu batzuen diametroa jaso da (cm):

45-48-56-34-47-38-52-44

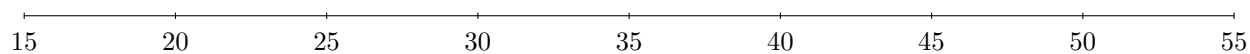
Datuetan jasotzen den informazio guztia kontuan hartzen duten neurriak kalkulatu, azter ezazu datu horietarako banakuntza normala egokia izan daitekeen. Datuak metroan izango balira, aldatuko al lirateke neurrien balioak?

Ebazpenak

1. ariketa

(a)

Puntu-diagrama:



Interpretazioa

(b) Datu guztiak ordenaturik:

$$16 - 19 - 20 - 21 - 21 - 24 - 24 - 26 - 28 - 29 - 30 - 32 - 36 - 42 - 50$$

Me =

 \bar{x} =

(c)

- \bar{x} -etik gora zenbat datu:
- \bar{x} -etik behera zenbat datu:

(d)

x_i	x_i^2	$(x_i - \bar{x})^3$
26	676	-6.504
24	576	-57.811
29	841	1.456
30	900	9.709
21	441	-323.771
19	361	-697.078
16	256	-1671.038
32	1024	70.616
36	1296	538.029
21	441	-323.771
20	400	-486.824
42	1764	2823.149
50	2500	10842.776
28	784	0.002
24	576	-57.811
418	12836	10661.13

 \bar{x} =

$$s_x = \sqrt{\frac{\sum_i x_i^2}{n} - \bar{x}^2} =$$

$$A_F = \frac{\sum_i (x_i - \bar{x})^3}{n s_x^3} =$$

Interpretazioa:

(e)

Bi datu handienak (42, 50) kenduta:

x_i	x_i^2	$(x_i - \bar{x})^3$
26	676	0.8
24	576	
29	841	
30	900	
21	441	
19	361	
16	256	
32	1024	
36	1296	
21	441	
20	400	
28	784	
24	576	
326	8572	599.4

$$\bar{x} =$$

$$s_x = \sqrt{\frac{\sum_i x_i^2}{n} - \bar{x}^2} =$$

$$A_F = \frac{\sum_i (x_i - \bar{x})^3}{\frac{n}{s_x^3}} =$$

Interpretazioa:

(f) Datuei simetriarako aldakuntza (erro karratua) aplikatuz:

x_i	$z_i = \sqrt{x_i}$	z_i^2	$(z_i - \bar{z})^3$
26	5.10		
24	4.90		
29	5.39		
30	5.48		
21	4.58		
19	4.36		
16	4.00		
32	5.66		
36	6.00		
21	4.58		
20	4.47		
42	6.48		
50	7.07		
28	5.29		
24	4.90		
418	78.25	418	5.6

$$\bar{z} =$$

$$s_z = \sqrt{\frac{\sum_i z_i^2}{n} - \bar{z}^2} =$$

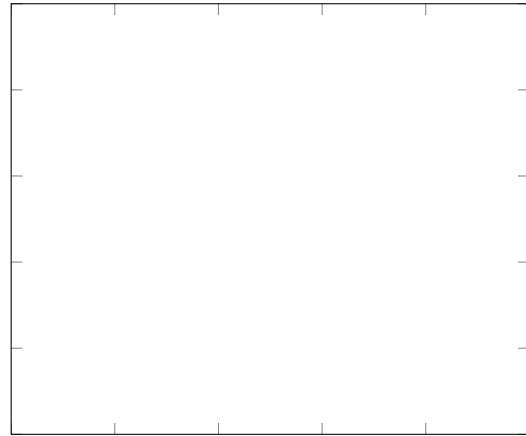
$$A_F = \frac{\sum_i (z_i - \bar{z})^3}{\frac{n}{s_z^3}} =$$

Interpretazioa:

2. ariketa

Mediana kalkulatzeko interpolazio-eskema:

x_i	n_i	$n_i x_i$	N_i (n_i metatuak)
250	126	$250 \times 126 =$	126
350	459		$126 + 459 =$



Kuantila	Zenbatgarren datua	Zein tartetan dagoen	Interpolazioa: x kalkulatu	Kuantilaren balioa
Q_1			$\frac{x}{100} = \frac{100}{100} \rightarrow x =$	
Q_3				
$P_{12.5}$				
$P_{37.5}$				
$P_{62.5}$				
$P_{87.5}$				

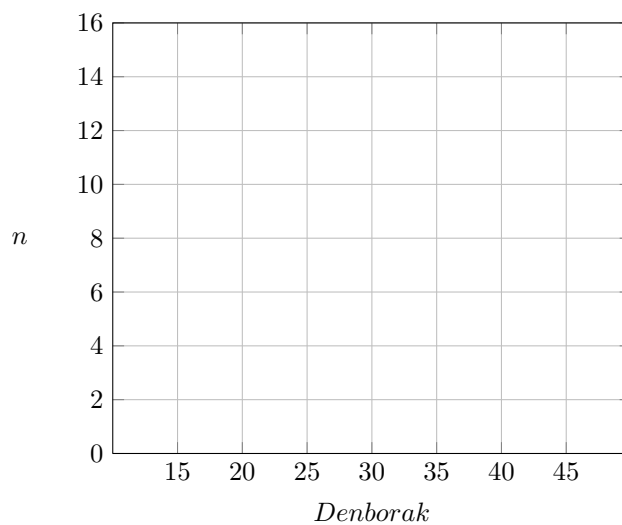
$$A_B = \frac{(Q_3 - Me) - (Me - Q_1)}{Q_3 - Q_1} =$$

$$K_M = \frac{(P_{87.5} - P_{62.5}) + (P_{37.5} - P_{12.5})}{P_{75} - P_{25}} =$$

3. ariketa

- Aldagai mota:
- Datu kopurua (handia/txikia):
- Beraz, grafikoa:

Denbora	Maiztasuna (n)
15-20	
20-25	
25-30	
30-35	
35-40	
40-45	



Datu ordenatuak:

18-19-20-20-21-22-22-23-24-24-25-25-25-25-26-26-26-26-27-27-27-27-28-29-29-32-32- 35-35-38-42

Kuantila	Zenbatgarren datua	Ondoko datuak	Formularen aplikazioa: $(1 - z) \times 1. \text{ datua} + z \times 2. \text{ datua}$
Me			
Q_1			
Q_3			
$P_{12.5}$			
$P_{37.5}$			
$P_{62.5}$			
$P_{87.5}$			

$$A_B = \frac{(Q_3 - Me) - (Me - Q_1)}{Q_3 - Q_1} =$$

$$K_M = \frac{(P_{87.5} - P_{62.5}) + (P_{37.5} - P_{12.5})}{P_{75} - P_{25}} =$$