Metodologia formulak

Aldagairen balioak X_i	Maiztasun absolutuak n_i	Maiztasun absolutu metatuak n_a	Maiztasun erlatiboak p_i	Maiztasun erlatiboa metatuak p_a	Maiztasun erlatibo portzentualak P_i	Maiztasun erlatibo portzentual metatuak P_a
X_1	n_1	n_1	n_1/N	n_{a1}/N	p_1*100	$p_{a1}*100$
X_2	n_2	n_1+n_2	n_2/N	n_{a2}/N	p_2*100	$p_{a2}*100$
...	...	...	...	...	...	...
X_n	n_n	N	n_n/N	1	p_n*100	100
	N		1		100	

M ABS guztien batura = M ABS METren azken zifra= subjektu kopuru totala (N)

ALDAGAI KUANTITATIBO JARRAITUAK EDO DATU TALDEKATUAK (TARTEKA KALKULUAK)

IBILTARTEA:

Xmax – Xmin + 2 * 0,5 * Neurgailuaren zehaztapena edo neurri-unitatea

Adibidea: 11 – 0 + 2 * 0,5*1 = 12

BENETAKO MUGAK edo MUGA ERREALAK: Agerizko muga ez da neurri erreala

X (itxurazko/ageriazko muga) ± 0,5* Neurgailuaren zehaztapena edo neurri unitatea

→ BEHE-MUGA ERREALA: X- (0,5 *1/0,1/0,01/...)
→ GOI-MUGA ERREALA: X + (0,5 *1/0,1/0,01/...)

Adibidea (0-3 tartea)

- BM: 0 - 0,5 * 1= - 0,5
- GM: 3 + 0,5 * 1 = 3,5

TARTE ZABALERA:

Goi-muga erreala – Behe-muga erreala

Adibidea: 0 – 3 tartea ➞Tarte zabalera: 3,5 – (– 0,5) = 4

TARTE-ORDEZKARIA (T.O.):

(Goi-muga + Behe-muga)/2

Adibidea: 0 – 3 tartea ➞Tarte-ordezkaria: (3 + 0)/2 = 1,5

Itxurazko mugak	Muga errealak	Tarte-ordezkaria
0-3	-0,5 - 3,5	1,5
4-7	3,5 - 7,5	5,5
8-11	7,5 - 11,5	9,5

BATUKARIAK (Σ)

Garrantzitsua i) eta j) desberdintzea

a) $\sum x_i = x_1 + x_2 + x_3 \ldots$

b) $\sum kX = k\sum X$ Konstantearen biderkadura (K * aldagaia)

c) $\sum k = nk$ Konstantearen batukaria (k*subjektu kopuru totala N)

d) $\sum(X+k) = \sum X + nk = (x,+k) + (x_2 + k) + (x_3 + k) \ldots$ Aldagai eta konstantearen batura

e) $\sum(X+Y) = \sum X + \sum Y$ Batukariaren banaketa

f) h)$\sum(X+Y)^2 = \sum X^2 + \sum Y^2 + 2\sum XY$ Binomioaren berretura

g) $\sum x \cdot \sum y = (x_1 + x_2 + x_3 \ldots) \cdot (y_1 + y_2 + y_3 \ldots)$

h) $\sum XY = X_1Y_1 + X_2Y_2 + \cdots + X_nY_n$ Aldagaien biderkaduren batukaria

i) $(\sum X)^2 = \sum X \sum X$ **Batukariaren berretura**

- X ALDAGAIAREN BATUKARIAREN KARRATUA

j) $\sum X^2 = X_1^2 + X_2^2 + \cdots + X_n^2$ **Balioen berreturen batukaria**

- X ALDAGAIAREN BALIOAREN KARRATUAREN BATUKARIA

X_i / Y_i: aldagai bakoitzaren balioa
i=1 ➞X edo Y aldagaiek hartzen duten 1. balioa
n: X edo Y aldagaiek hartzen duten azken balioa
edo aldagaiaren balio kopurua
k: konstantea

JOERA ZENTRALEKO NEURRIAK

BATAZ BESTEKO ARITMETIKOA ($\bar{X}$)

- **DATU ISOLATUEKIN:** : n balioen batura datu kopuruaz zatituta
  - $\bar{X} = \frac{\sum X_i}{N} = \frac{2+3+5+10}{4} = 5$ (Xi: 2,3,5,10)

- **DATUAK MAIZTASUN-TAULAN:** aldagaiaren balioa eta bere maiztasun absolutuaren arteko biderketen batukaria, datu kopuruaz zatituta
  - $\bar{X} = \frac{\sum X_i \cdot n_i}{N} = \frac{30}{12} = 2{,}5$

X_i	n_i	$X_i.n_i$
5	1	5
3	5	15
2	4	8
1	2	2
	N= 12	30

- **MAIZTASUN- TAULAN DATU TALDEKATUEKIN:** (Xi ➞ T.O) tarte ordezkarien eta bere maiztasun absolutuaren arteko biderketen batukaria, datu kopuruaz zatituta
  - $\bar{X} = \frac{\sum X_i(T.O.) \cdot n_i}{N} = \frac{430}{20} = 21{,}5$
  - 💡 T.O. = (Goi-muga + Behe-muga)/2

X_i	n_i	T.O.	$X_i.n_i$
0-9	2	4,5	9
10-19	5	14,5	72,5
20-29	10	24,5	245
30-39	3	34,5	103,5
	N=20		

BATEZ BESTEKOAREN EZAUGARRIAK:

1. Lagin batean elementu guztien batez besteko aritmetikoarekiko desbiderazioen baturak 0 balio du.
   → $\sum(Xi-\bar{X})=0$
2. Aldagai bati konstante bat gehitu, kendu, biderkatu edo zatitzen zaionean, batez bestekoari gauza bera gertatzen zaio

$Y_i = X_i + k$

$\bar{Y} = \bar{X} + k = 5 + 1 = 6$

$\bar{X} = \frac{\sum X_i}{N} = \frac{20}{4} = 5$

$\bar{Y} = \frac{\sum Y_i}{N} = \frac{24}{4} = 6$

TALDE OSOAREN BATAZ BETESKOA

Kasu batzuetan, lagin-baten barnean azpitaldeak egon. Azpitalde bakoitzaren media badaukagu, lagin osoaren media (talde osoaren bataz bestekoa) kalkula dezakegu .

→ datu-multzo bat azpimultzotan banatuta dagoenean, multzo horien bataz besteko aritmetikoen batez besteko ponderatua

- 1. taldea: 2, 3, 4, 5, 6 (n=5) ➞ batez bestekoa= 4
- 2. taldea: 2, 3, 4, 5, 6, 7, 8 (n=7) ➞ batez bestekoa= 5

$$\overline{X_o} = \frac{\overline{X_1} \cdot n_1 + \overline{X_2} \cdot n_2}{n_1 + n_2} = \frac{(4 \cdot 5) + (5 \cdot 7)}{5 + 7} = 4,58$$

BATAZ BESTEKO PONDERATUA

Elementu guztien "pisua" edo garrantzia desberdina denean erabiltzen da. Kontuan hartu behar da elementu guztiek ez dutela pisu bera.

- p_i = pisua

$$\overline{X_p} = \frac{\sum X_i \cdot p_i}{\sum p_i} = \frac{(5 \cdot 3) + (8 \cdot 1) + (7 \cdot 2) + (5 \cdot 4)}{3 + 1 + 2 + 4} = 5,7$$

	p_i	X_i
Lehen partziala	3	5
Talde lana	1	8
Lan indibiduala	2	7
2. Partziala	4	5

MEDIANA (ME)

- Muga itxi gabeak
- Datu sakabanatuak
- Banaketa asimetrikoa
- Aldagaiak gutxienez ordinalak

2 pauso:

1. Erdiko subjektuaren posizioa identifikatzea (adib. 4. postuan)
2. zein da medianaren balioa? Posizio horretako subjektuaren balioa

DATU ISOLATUAK

Kontuan hartu behar da laginaren tamaina edo subjektu kopurua (bakoitia/bikoitia)

- Subjektu kop **bakoitia** ➞ erdian subjektu bakarra
  1. $(N+1)/2 = (9+1)/2 = 5$ posizioa
  2. 5 posizioko puntuazioa ➞ MEDIANA **(Me = 7)**

	1.	2.	3.	4.	**5.**	6.	7.	8.	9.
X_i	1	3	3	5	**7**	8	9	9	9

- Subjektu kop **bikoitia** ➞ erdian 2 subjektu (mediana = 2 subjektuen bataz bestekoa)
  1. $(N+1)/2 = (8+1)/2 = 4,5$ posizioa
  2. 4 eta 5 posizioen puntuazioa ➞ 3 eta 7; horien media ➞ **[Me = (3+7)/2= 5]**

	1.	2.	3.	**4.**	**5.**	6.	7.	8.
X_i	1	2	3	**3**	**7**	8	9	10

DATU ISOLATUAK MAIZTASUN-TAULAN

Medianaren balioa erraztu egiten da maiztasun-taulen bidez. Garrantzia handia m abs metatatuak.

1. Erdiko subjektuaren posizioa $(N+1) / 2 = (11+1) / 2 = 6$
2. Posizio horretako subjektuaren balio bilatu, horretarako m abs met-en (na) 6/+ den m met bilatu ➞ na horri dagokion puntuazioa (Xi) ➞ mediana **(Me = 5)**

X_i	n_i (m abs)	n_a (m abs met)
2	1	1
3	4	5
5	**3**	**8**
10	3	11
	11	

DATU TALDEKATUAK MAIZTASUN-TAULAN (TARTEAK)

1. Medianadun tartearen poszioa: $(N+1)/2 = (15+1)/2 = 8$
2. Iortutako balioa n_a-n aurkitu ➞ 8/+ den aurreneko n_a **10** denez 85-99 tartean dago mediana
3. Formula tarte horretako datuak erabiliz aplikatu:

$$M_e = L_i + \frac{\frac{N}{2} - n_{a-1}}{n_i} \cdot I_i = 84,5 + \left(\frac{\frac{15}{2} - 2}{8}\right) \cdot 15 = 94,81$$ ➞ non

X_i	n_i	n_a
70-84	2	2
85-99	8	10
100-114	4	14
115-129	1	15

L_i: Medianadun tartearen behe-muga erreala
I_i: Medianadun tartearen tarte-zabalera
→ = goi-muga erreala – behe-muga erreala
n_i: Medianadun tartearen m abs
n_{a-1}: Medianadun tartearen aurreko m met

MODA

- **Definizioa**: lagin batean gehien errepikatzen den puntuazioa.
  - Zelan identifikatu? ➞ Maiztasun tauletan, m abs (n_i) handiena duen balioa

Xi	ni
Gipuzkoa	4
Araba	**6**
Nafarroa	3
Bizkaia	4

- **Erabilera**: edozein aldagai motarekin erabil daiteke, baita aldagai kualitatiboekin ere
- Gerta daiteke hainbat aldagaiek balio handiena izatea, horregatik, datuen arabera:
  - Moda bakarra **(moda bakarreko banaketa)**
  - Bi moda **(banaketa bimodala)**
  - Hainbat moda (banaketa multimodala)

Adibidea: Ikasleen jatorria (G= Gipuzkoa, B= Bizkaia, A= Araba, N= Nafarroa)
G,G,G,G,A,A,A,A,A,A,N,N,N,B,B,B,B ➞Moda = Araba

BANAKAKO POSIZIO NEURRIAK

KONTZEPTUA: subjektu baten posizioa/ordena talde baten barruan.

→ KUARTILAK (Q): banaketa 4 ataletan (Q1=P25 ; Q2=P50 ; etab.)
→ DEZILAK (D): banaketa 10 ataletan (D1=P10 ; D2=P20 ; etab.)
→ PERTZENTILAK (P): banaketa 100 ataletan (ehunekoak bezala) KALKULUA: - Datu isolatuekin: hiruko erregela

DATU ISOLATUAK: Hiruko erregela

1. 7 subjektuen datuak txikitik handienera ordenatu ➞ Xi: 3, 3, **4**, 4, 5, 6, 7
2. Inkognitak (Xi edo P) hiruko erregelen bidez kalkulatu

a) 5 atera duen subjektuak zein subjektu portzentai uzten du bere azpian (P_k)? %57,14

→ *5 puntuaziotik bera 4 subjektu geratu, 100 subjektu izango bagenitu zenbat?*

7 ---------4
100 ----- P_k

Pk = (100 x 4) / 7 = 57,14

b) Zein puntuazio uzten du bere azpian sujektuen %40 (P40)? 3. Posizizoko balioa ➞ $P_{40} = 4$

100 subjetutatik 40 azpian geratzen badira, 7 subjektutatik zenbat?

100 -----40
7 ------X

Xi = (7 x 40) / 100 = 2,8 ≈ 3. posizioan dagoen balioa

MAIZTASUN-TAULAN: hiruko erregela + na

a) **$P_k = 4$; K? Hiruko erregela**
4 balioa baino gutxiago zenbat subjektu? Taulan: 9

X_i	n_i	n_a
2	4	4
3	5	**9**
4	6	15
5	7	22

22 —---9
100----$P_{k?}$

K= (100x9) /22 = 40,9

b) P_{25}? **$P_{25} = 3$**

100 —-25
22-----X

P_k = (N.k) / 100 = (22.25) / 100 = 5,5

DATU TALDEKATUAK

- $k = \frac{(P_k - L_i)\cdot\left(\frac{n_i}{I_i}\right) + n_{a-1}}{N} \cdot 100$
- $P_k = L_i + \frac{\frac{N\cdot k}{100} - n_{a-1}}{n_i} \cdot I_i$

K: Pertzentilen ordena edo maila
L_i: Pertzentila kokatzen den tartearen behe-muga erreala
I: Tarte-zabalera
N_{a-1}: Aurreko tarteraino metaturiko maiztasuna

Adibidea ➞ Kalkula ezazu 90. Pertzentila P_{90}=0?

1. **PERTZENTILAREN POSIZIOA:** $N \cdot k$ 100 (hiruko erregela)➞ x (posizioa) lortu

100----90
79------X_i

(N.K)/100 = (79.90)/100 =71

Xi	ni	na
20-24	22	22
25-29	25	47
30-34	**32**	**79**

2. Zein na-n (M ABS MET)? x/+ ➞ tartea identifikatu (balioa) (30-34)
3. **Ii:** tarte zabalera= goi-muga erreala – behe-muga erreala = 34,5-29,5 = 5
4. Behin behar ditugun datuak lortuta, **formularen bidez kalkulua egin**

$$P_k = L_i + \frac{\frac{N\cdot k}{100} - n_{a-1}}{n_i} \cdot I_i = P_{90} = 29{,}5 + \frac{\frac{79\cdot 90}{100} - 47}{32} \cdot 5 = \mathbf{33{,}27}$$

Zer portzentaia/ehuneko uzten du bere azpitik 25 puntu lortu zituen pertsona batek?

P_K=25 balioa (ez posizioa) barne hartzen duen tarteko datuak erabiliz aplikatu behar dugu formula

$$k = \frac{(P_k - L_i)\cdot\left(\frac{n_i}{I_i}\right) + n_{a-1}}{N} \cdot 100 = \frac{(25 - 24{,}5)\cdot\left(\frac{25}{5}\right) + 22}{79} \cdot 100 = 31$$

ALDAKORTASUN/SAKABANATZE NEURRIAK

Datu isolatuekin

$$S_x = \sqrt{\frac{\sum(X_i - \bar{X})^2}{N}}$$

Datuak maiztasun taula batean badaude (datu isolatuak)

$$S_x = \sqrt{\frac{\sum(X_i - \bar{X})^2 \cdot n_i}{N}}$$

Datuak maiztasun taula batean badaude (datu taldekatuak)

$$S_x = \sqrt{\frac{\sum(X_i[T.O.] - \bar{X})^2 \cdot n_i}{N}}$$

MAIZTASUN-TAULA SINPLEA

1. bataz besteko ateratzea (formulan beharrezkoa baita)
2. Taula osatzea falta diren formulako datuekin
3. Datu guztiak ditugunean formula jarraitu

X_i	n_i	$X_i.n_i$	$(X_i - \bar{X})^2\, n_i$
1	2	2	4,5
2	4	8	1
3	5	15	1,25
5	1	5	6,25
	N= 12	**30**	**13**

$$\bar{X} = \frac{\sum X_i \cdot n_i}{N} = \frac{30}{12} = 2{,}5$$

$$S_x = \sqrt{\frac{\sum(X_i - \bar{X})^2 \cdot n_i}{N}} = \sqrt{\frac{13}{12}} = 1{,}03$$

$$S_x^2 = \frac{13}{12} = 1{,}08$$

MAIZTASUN-TAULA DATU TALDEKATUEKIN

$\bar{X} = \frac{\sum X_i(T.O.)\cdot n_i}{N} = \frac{430}{20}$=21,5

$S_x = \sqrt{\frac{\sum(X_i[T.O.] - \bar{X})^2 \cdot n_i}{N}} = \sqrt{\frac{1420}{20}} = \sqrt{71} = 8{,}43$

Xi-ren ordez ➞ T.O. = (Goi-muga + Behe-muga)/2

X_i	n_i	T.O.	T.O. n_i	$(T.O. - \bar{X})^2 \cdot n_i$
0-9	2	4,5	9	578
10-19	5	14,5	72,5	245
20-29	10	24,5	245	90
30-39	3	34,5	103,5	507
	N=20			**1420**

TALDE OSOAREN BARIANTZA

Noiz? Azpitaldeak daudenean non

$$S_o^2 = \frac{\sum n_j \cdot S_j^2}{\sum n_j} + \frac{\sum n_j \cdot (\bar{X}_j - \bar{X}_o)^2}{\sum n_j}$$

N azpitalde tamaina
J azpitaldea
ENJ lagin totala

Adibidea:

A taldea: 7, 6, 4, 5, 9, 5, 3, 8, 8, 5 ➞ $\overline{X_A} = 6$ / $S_A = 1{,}84$

B taldea: 30, 40, 50, 50, 50, 60, 70, 80, 80, 90 ➞ $\overline{X_B} = 60$ / $S_B = 18{,}4$

1. $\overline{X_o} = \frac{\overline{X_1}\cdot n_1 + \overline{X_2}\cdot n_2}{n_1 + n_2} = \frac{(6\cdot 10)+(60\cdot 10)}{10+10} = 33$
2.

$$S_o^2 = \frac{\sum n_j \cdot S_j^2}{\sum n_j} + \frac{\sum n_j \cdot (\bar{X}_j - \overline{X_o})^2}{\sum n_j}$$

$$= \frac{(10 \cdot 1{,}84^2) + (10 \cdot 18{,}4^2)}{10+10} + \frac{10 \cdot (6-33)^2 + 10 \cdot (60-33)^2}{10+10}$$

$$= \frac{33{,}86 + 3385{,}6}{20} + \frac{7290 + 729}{20} = 170{,}97 + 400{,}95 = 571{,}92 \qquad S_o = 23{,}91$$

ALDAKUNTZA KOEFIZIENTEA

Bi aldagaien sakabanatze-maila konparatzeko

$$AK = \frac{S_x}{\bar{X}} \cdot 100$$

KOARTIL ARTEKO IBILTARTEA (Q)

Bataz bestekoa ez da beti neurri egokia, desegokia izanez gero, desbideratze tipikoa desegokia da

- Banaketa asimetrikoa + Muturretan balio arraroak

$$Q = \frac{Q_3 - Q_1}{2} = \frac{P_{75} - P_{25}}{2} \qquad P_k = L_i + \frac{\frac{N\cdot k}{100} - n_{a-1}}{n_i} \cdot I_i$$

X_i	n_i	n_a
1-3	15	15
4-6	20	35
7-9	10	45
10-12	5	50

ADIBIDEA:

P_{75}ren posizioa $\frac{N\cdot K}{100} = \frac{50\cdot 75}{100} = 37{,}5$ ➞ P_{75}ren kalkulua $P_{75} = 6{,}5 + \frac{37{,}5-35}{10} \cdot 3 = 7{,}25$

P_{25}ren posizioa $\frac{N\cdot K}{100} = \frac{50\cdot 25}{100} = 12{,}5$ ➞ P_{25}ren kalkulua $P_{25} = 0{,}5 + \frac{12{,}5-0}{15} \cdot 3 = 3$

$$Q = \frac{Q_3 - Q_1}{2} = \frac{P_{75} - P_{25}}{2} = \frac{7{,}25 - 3}{2} = 2{,}12$$

BARIANTZA ETA DESBIDERATZE TIPIKOAREN EZAUGARRIAK

a)Balio positiboak: S ≥ 0 eta S 2 ≥ 0

b) Aldagai bati konstante bat gehitzen badiogu, bere bariantza eta desbideratze tipikoa ez dira aldatzen

c) Aldagai bat bider konstante bat egiten badugu, Sx konstantearen balioagatik biderkatua geratuko da, eta S 2 x konstantearen berreduragatik.

FORMARI BURUZKO NEURRIAK

ASIMETRIA

FISHER-EN ASIMETRIA INDIZEA

Maiztasun-taulan: (tarterik gabe)

$$a_3 = \frac{\sum n_i (X_i - \bar{X})^3 / N}{S_X^3}$$

Maizt-taul (tarterekin): Xi-ren ordez T.O.

Datu isolatuak: ni kendu ➞ **Balio isolatuetan n_i=1**, beraz, formulan kendu dezakegu

→ Aurreikusten badugu ez dela simetrikoa izango alternatiba : alborapen indize koartilikoa?

ALBORAPEN INDIZE KOARTILIKOA

$$\boldsymbol{a_q = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_2) + (Q_2 - Q_1)}}$$

→ Emaitza beti: (-1, 1) ➞ $a_{q=}$ +1 eta –1 bitartean

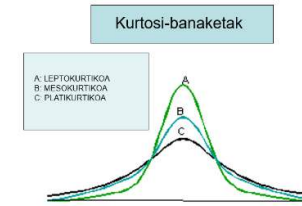
INTERPRETAZIOA

- **Asimetria indizea = 0** ➞ Banaketa simetrikoa
- **Asimetria indizea = positiboa** ➞ Asimetria positiboa
- **Asimetria indizea = negatiboa** ➞ Asimetria negatiboa

ZORROZTASUN NEURRIAK (KURTOSIA)

FISHER-EN KURTOSI INDIZEA (a_4)

$$a_4 = \frac{\sum n_i (X_i - \bar{X})^4 / N}{S_X^4} - 3$$



- Balio isolatuetan n_i=1, beraz, formulan kendu dezakegu
- Maizt-taul (tarterekin): Xi-ren ordez T.O.

INTERPRETAZIOA

MESOKURTIKOA (Zorroztasun indizea = 0): Banaketa normala

LEPTOKURTIKOA(Zorroztasun indizea = positiboa): Banaketa normala baino zorroztasun handiagoa

PLATIKURTIKOA(Zorroztasun indizea = negatiboa): Banaketa normala baino zorroztasun txikiagoa

ZORROZTASUN INDIZE PERTZENTILIKOA

Alternatiba bataz-bestekoa desegokia bada:

$$K_p = \frac{(Q_3 - Q_1)/2}{P_{90} - P_{10}} = \frac{(P_{75} - P_{25})/2}{P_{90} - P_{10}}$$

INTERPRETAZIOA

- Kp=0,263 B.normala
- Kp<0,263 Platikurtikoa
- Kp>0,263 Leptokurtikoa

PUNTUAZIO TIPIKOAK ETA ERATORRIAK

- Puntuazio zuzenak Xi
- Puntuazio diferentzialak $Xi - \bar{X}$
- Puntuazio tipikoak (z) Z = $Xi - \bar{X}$ Sx

Puntuazio eratorria= S_t x Z + $\overline{X_t}$

💡 $\overline{X_t} = bataz\ bestekoa$
$S_t = desbideratze\ tipikoa$

Eskala eratorriak

■ T = 10 x Z + 50 (T puntuazioak: $\bar{\bar{Xt}}$ = 50 eta St=10)

T = 10 x Z + 50

■ S = 2 x Z + 5 (Eneatipo edo Estaninoak: $\bar{\bar{Xs}}$ = 5 eta Ss=2)

S = 2 x Z + 5

■ CI = 15 x Z + 100 (Adimen-kozientea Wais: $\bar{\bar{Xci}}$ = 100 eta Sci=15)

CI = 15 x Z + 100