

ESTADÍSTICA ETA DATUEN ANALISIA

Azterketa ebatziak.

2018-2019 ikasturtea

Donostiako Ekonomia eta Enpresa Fakultatea. EHU

Egilea eta irakasgaiaren irakaslea: Josemari Sarasola



Gizapedia

gizapedia.hirusta.io

ESTADISTIKA ETA DATUEN ANALISIA

Irakaslea: Josemari Sarasola

Data: 2019ko urtarrilaren 8a, 16:00

Iraupena: Ordu t'erdi.

Praktika: Adierazburu-orria

I. ebazkizuna: pobreziaren neurketa (2.5 puntu)

Herri batean familien lagin bat jaso eta horien tamaina (heldu eta haur kopurua) eta errenta jaso dira:

Kideak	2H, 2h	2H	2H, 1h	1H	3H	2H, 2h
Errenta (€)	2200	2000	1500	900	1800	2000

Pobrezia-atalasea 800€ da, baliokidetasun-eskala aplikatu ondoren.

Egin beharreko atazak:

- Kalkulatu pobrezia-tasa eta interpretatu hitzez.
- Pobreziaren intentsitatea kalkulatu ezazu.
- Pobreziaren adierazle sintetikoa eman ezazu. Zer adierazten du?

Oharra: kalkuluetarako OCDEko baliokidetasun-eskala aplikatu.

II. ebazkizuna: asoziazioa (2 puntu)

Eskualde batean 90 urtetik gorako pertsonen artean sexua eta Alzheimer gaixotasuna (AG) duten jaso dira. Datuak kontingentzia-taula honetan bildu dira:

Sexua(↓)/AG? →	Ez	Bai
Gizona	87	92
Emakumea	157	98

Egin beharreko atazak:

- Kalkulatu khi-karratua oinarrituriko neurriren bat eta interpretatu ezazu.
- Kalkulatu lambda eta interpretatu ezazu.

III. ebazkizuna: zenbaki indizeak (2 puntu)Lantegi batean zura eta burdina baliatzen dira. 2016-2018 urteetan kontsumitutako q kopuruak (tonak) eta p prezioak (€ tonako) jaso dira:

Urtea	p(zura)	p(burdina)	q(zura)	q(burdina)
2016	10	30	5	15
2017	14	42	6	10
2018	19	40	8	13

Egin beharreko atazak:

- Kalkulatu Laspeyres prezio-indizeak 2016 oinarritzat harturik. Indize horietan oinarriturik, zenbat igo ziren prezioak 2018an 2017 urteari buruz? Prezio-igoera hori zehatza al da? Zergatik?
- Kalkulatu zuraren eta burdinaren erreperkutsioak eta partaidetzak 2017 urterako, eta interpretatu itzazu.

IV. ebazkizuna: egunkari saltzailearen modelo (2 puntu)

Dendari batek ordenagailuak erosi behar ditu. Ordenagailu horien eskaria hurrengo hilabetean honela banatzen dela uste du:

Ordenagailuak	4	5	6	7	8	9
Probabilitatea	0.15	0.2	0.25	0.25	0.1	0.05

Ordenagailuak 100€-na kostatu zaizkio, eta 250€-tan saltzen ditu. Hurrengo hilabetean saltzen ez dituenak 50€-tan likidatu ditu.

Egin beharreko atazak:

- Mozkina maximotzeko erosi beharreko ordenagailu kopurua kalkulatu behar da.
- Kopuru optimo horretarako, galduko diren salmenten batezbestekoa, unitatetan.
- Kopuru optimo horretarako, likidatzeko geratuko diren unitateen batezbestekoa.

ESTATISTIKA ETA DATUEN ANALISIA

Irakaslea: Josemari Sarasola

Data: 2019ko urtarrilaren 8a, 16:00

Iraupena: Ordu t'erdi.

Praktika: Ebazpena

I. EBAZKIZUNA

Lehenbizi, baliokidetasun-eskala aplikatu behar zaie familiei:

Kideak	2H, 2h	2H	2H, 1h	1H	3H	2H, 2h
Kide baliokideak (1+0.7H+0.5h)	2.7	1.7	2.2	1	2.4	2.7
Errenta baliokideak	814 (4)	1176 (2)	681 (3)	900 (1)	750 (3)	740 (4)

Lagineko 17 banakoen errenta baliokideak ordenaturik zerrendatuz:

681-681-681-740-740-740-740-750-750-750-814-814-814-814-900-1176-1176

(a) Pobrezia-atalasea 800€ denez, **pobrezia-tasa** $10/17=0.58$, hots, pobreak laginaren %58 dira.

(b)

$$I = \frac{\sum_{i=1}^p (z - x_i)}{pz} = \frac{(800 - 681) \times 3 + (800 - 740) \times 4 + (800 - 750) \times 3}{10 \times 800} = 0.093$$

Pobreziaren intentsitateak 0-1 bitarteko balioak hartzen ditu, eta 0tik zenbat eta gertuago, orduan eta intentsitate txikiagoa. $I = 0.093$ izanda, printzipioz ezin esan pobreziaren intentsitatea handia denik.

(c)

Pobreziaren adierazle sintetikoa kalkulatu behar da:

$$S = \frac{2 \sum_{i=1}^p (z - x_i)(p + 1 - i)}{(p + 1)nz} = \frac{2[(800 - 681) \times 10 + (800 - 681) \times 9 + (800 - 681) \times 8 + (800 - 740) \times 7 + (800 - 740) \times 6 + (800 - 740) \times 5 + (800 - 740) \times 4 + (800 - 750) \times 3 + (800 - 750) \times 2 + (800 - 750) \times 1]}{(10 + 1) \times 17 \times 800} = 0.064$$

Adierazle sintetikoak pobre kopurua nahiz pobreziaren intentsitatea bateratzen dituen adierazle sintetiko edo integrala da. Ietik zenbat eta gertuago, pobrezia orduan eta handiagoa da. Ezin esan, beraz, pobrezia handia denik, lagin errearen eta antzeko ikerketen erreserbapean betiere.

II. EBAZKIZUNA

(a): Khi-karratu

Bazter maiztasunak eta lagin tamaina kalkulatu ditugu, O maiztasun empirikoak ere jarrita:

87	92	179
157	98	255
244	190	434

Maiztasun teorikoak (E) kalkulatu ditugu:

$\frac{179 \times 244}{434} = 100.63$	78.36
143.36	111.63

 $(O - E)^2/E$ kalkulatu ditugu, eta horien erroa (parentesien artean), dagokien zeinuarekin:

$\frac{(87-100.63)^2}{100.63} = 1.84$ (-1.35)	2.73 (1.65)
1.49 (1.22)	1.66 (-1.28)

Khi-karratu kalkulatu dugu:

$$\mathbf{X^2} = 1.84 + 2.73 + 1.49 + 1.66 = 7.72$$

Khi-karratua oinarrituriko neurri sinpleena, 2×2 tauletan, phi da:

$$\phi = \sqrt{\frac{7.72}{434}} = 0.13$$

Asoziazioa bi aldagaien artean ahula da beraz (< 0.2), lagin errorearen eta antzeko ikerketen erreserbapean.Ahula izanda ere, norabidea aztertzen bada, gizonen Alzheimer gaixotasuna ez izateko joera EZA nabarmenena dela esango genuke, $(O - E)/\sqrt{E}$ balioei erreparatuz (parentesien artean).

(b): Lambdaren kalkulua

- Aldagai independentea: sexua
- Aldagai dependentea: gaixotasuna

Aldagai independentea, sexua, kontuan hartu gabe, auresana da gaixotasuna ez dela garatuko ($244 > 190$), baina horrela 190 errore egiten dira.

Sexua kontuan harturik berriz,

- Gizona bada, auresana da Bai ($92 > 87$) eta erroreak 87
- Emakumea bada, auresana da Ez ($157 > 98$) eta erroreak 98

Sexua kontuan harturik beraz, guztira $87+98=185$ errore, sexua kontuan hartu gabe baino $190-185=5$ errore gutxiago.

Beraz,

$$\lambda = \frac{5}{190} = 0.02$$

Eta beraz, asoziazioa ahula dela esan daiteke (< 0.2), lagin errorearen eta antzeko ikerketen erreserbapean.

III. EBAZKIZUNA

(a)

$$L_{p,2016}^{2017} = \frac{14 \times 5 + 42 \times 15}{10 \times 5 + 30 \times 15} \times 100 = 140$$

$$L_{p,2016}^{2018} = \frac{19 \times 5 + 40 \times 15}{10 \times 5 + 30 \times 15} \times 100 = 139$$

Prezioak igoera kalkulatzeko, 2017-2018 bitartean:

$$(139/140) \times 100 = 99.28$$

Eta hortik prezio igoera $99.28 - 100 = -0.72\%$ izan zen, 0.72% ko jaitsiera beraz.

Emaitza hori ez da guztiz zehatza, Laspeyres indizea zirkularra ez delako, Aldaketa zehatza kalkulatzeko, 2017 urtea hartu genuke oinarritzat, eta kalkuluak berregin 2018 urterako.

(b)

Erreperkutsioak kalkulatzeko:

$$R_{zura,2016}^{2017} = \frac{(14 - 10) \times 5}{10 \times 5 + 30 \times 15} = 0.04$$

$$R_{burdina,2016}^{2017} = \frac{(42 - 30) \times 15}{10 \times 5 + 30 \times 15} = 0.36$$

2017 urtean izandako prezioen 40% ko igoeratik 4 puntu zurari dagozkio, eta 36 burdinari.

Partaidetzak erreperkutsio portzentualak besterik ez dira:

$$P_{zura,2016}^{2017} = \frac{36}{40} \times 100 = 90$$

$$P_{burdina,2016}^{2017} = \frac{4}{40} \times 100 = 10$$

Hartara, 2017 urtean izandako prezioen igoeratik 10% zurari dagokio, eta 90% burdinari.

IV. EBAZKIZUNA

(a)

Ordenagailu kopuru ezberdinetarako itxarondako mozkina kalkulatzeko duguz:

4 ordenagailu erosita,

Salmentak	Mozkina (x)	p(x)	xp(x)
4	$4 \times 150 = 600$	1	600
		1	600

5 ordenagailu erosita,

Salmentak	Mozkina (x)	p(x)	xp(x)
4	$4 \times 150 - 1 \times 50 = 550$	0.15	82.5
5	$5 \times 150 = 750$	0.85	637.5
		1	720

6 ordenagailu erosita,

Salmentak	Mozkina (x)	p(x)	xp(x)
4	$4 \times 150 - 2 \times 50 = 500$	0.15	75
5	$5 \times 150 - 1 \times 50 = 700$	0.20	140
6	$6 \times 150 = 900$	0.65	585
		1	800

7 ordenagailu erosita,

Salmentak	Mozkina (x)	p(x)	xp(x)
4	$4 \times 150 - 3 \times 50 = 450$	0.15	67.5
5	$5 \times 150 - 2 \times 50 = 650$	0.20	130
6	$6 \times 150 - 1 \times 50 = 850$	0.25	212.5
7	$7 \times 150 = 1050$	0.40	420
		1	830

8 ordenagailu erosita,

Salmentak	Mozkina (x)	p(x)	xp(x)
4	$4 \times 150 - 4 \times 50 = 400$	0.15	60
5	$5 \times 150 - 3 \times 50 = 600$	0.20	120
6	$6 \times 150 - 2 \times 50 = 800$	0.25	200
7	$7 \times 150 - 1 \times 50 = 1000$	0.25	250
8	$8 \times 150 = 1200$	0.15	180
		1	810

Itxarondako mozkin handiena 7 ordenagailu erosita gertatzen da. Beraz, 7 ordenagailu erosi behar dira.

(b)

7 ordenagailu erosita,

Eskaria	Galtzen diren salmentak (x)	p(x)	xp(x)
≤ 7	0	0.85	0
8,9	1	0.15	0.15
		1	0.15

Beraz, batezbeste 0.15 ordenagailuren salmenta galduko da.

(c)

7 ordenagailu erosita,

Eskaria	Likidatzeko geratuko direnak (x)	p(x)	xp(x)
4	3	0.15	0.45
5	2	0.2	0.4
6	1	0.25	0.25
≥ 7	0	0.40	0
		1	1.1

Beraz, batezbeste 1.1 ordenagailu gertauko dira likidatzeko.

ESTATISTIKA ETA DATUEN ANALISIA

Irakaslea: Josemari Sarasola

Data: 2018ko URTARRILAREN 8a

Iraupena: 35 minutu

Erantzun bakarra da zuzena galdera bakoitzean. Guztira testak 2 puntu balio du. Erantzun zuzenak 0.1 puntu balio du. Erantzun oker bakoitzak horren erdia kentzen du. Galderak erantzun gabe utz daitezke, punturik gehitu eta kendu gabe.

- (1) Honako neurri hauetatik, zein da sendoa edo jasankorra?
 - (a) Batezbesteko aritmetiko sinplea
 - (b) Kuartil arteko ibiltartea
 - (c) Desbideratze estandarra
 - (d) Bat ere ez
- (2) Emakumeen batez besteko kalifikazioa 8 da; gizonena, 7. Batez besteko kalifikazioetan sexu diferentziaren indarra neurtzeko, zer kalkula daiteke?
 - (a) Eta
 - (b) Phi
 - (c) Pearson korrelazio koefiziente lineala
 - (d) Aurreko guztiak dira zuzenak
- (3) Irakasgai batean 10 bertako ikasle eta 8 atzerritar daude. Zein grafiko da egokiena bi ikasle talde horien kalifikazioak (0, 0.1, ..., 9.9, 10) alderatzeko?
 - (a) Back-to-back edo ondoko histogramak
 - (b) Barra diagramak
 - (c) Puntu diagramak
 - (d) Maiztasun poligonoak
- (4) Zein da desbideratze estandarren definizio zuzena?
 - (a) Datu bakoitza batezbesteko aritmetikotik batezbestez zenbat desbideratzen den adierazten du, horretarako desbideratzeen batezbesteko kuadratikoa erabiliz.
 - (b) Datu bakoitza beste guztietatik batezbestez zenbat desbideratzen den adierazten du, horretarako desbideratzeen batezbesteko aritmetikoa erabiliz.
 - (c) Datu bakoitza hurrengo datutik batezbestez zenbat desbideratzen den adierazten du, horretarako desbideratzeen batezbesteko aritmetikoa erabiliz.
 - (d) Datu bakoitza batezbesteko aritmetikotik batezbestez zenbat desbideratzen den adierazten du, horretarako desbideratzeen batezbesteko aritmetikoa erabiliz.
- (5) Zein da Cohenen d formula erabiltzeko baldintza?
 - (a) Bi datu multzoek datu kopuru berdinak eduki behar dituzte.
 - (b) Bi datu multzoek bariantza berdinak eduki behar dituzte, lagin-errorea gorabehera.
 - (c) Bi datu-multzoak populazio beretik eratorriak izan behar dira.
 - (d) Bi datu-multzoak lagin berari buruzkoak izan behar dira.
- (6) Zertarako erabiltzen da Cohenen d neurria?
 - (a) Bi bariantza zenbateraino diren desberdinak neurtzeko.
 - (b) Bi datu multzoen datu estandartuen batezbestekoen diferentzia neurtzeko.
 - (c) Bi datu multzoen baterako sakabanatzea neurtzeko.
 - (d) Bi batezbesteko zenbateraino diren desberdinak neurtzeko.

Honako datu hauek hurrengo lau galderetarako erabiltzen dira:

x	4	0
y	0	2

- (7) Zenbat da x aldagaiaren desbideratze estandarra (zuzendu gabea)?
- 1
 - 2
 - 3
 - 4
- (8) Zenbat da y aldagaiaren kuasibariantza edo bariantza zuzendua?
- 1
 - 2
 - 3
 - 4
- (9) Zenbat da x eta y aldagaien arteko kobariantza?
- 1
 - 2
 - 1
 - 2
- (10) Pearson korrelazio koefiziente lineala kalkulatu, nola interpretatu behar da?
- Korrelazioa alderantzizkoa eta perfektua da.
 - Korrelazioa alderantzizkoa eta sendoa da.
 - Korrelazioa zuzena eta perfektua da.
 - Korrelazioa zuzena eta sendoa da.
- (11) Zertarako erabiltzen da kuasibariantza edo bariantza zuzendua?
- Bariantza datu atipiko edo muturrekoen eraginetik aparte mantentzeko.
 - Datuak tartetan izanik, erdipuntua hartzeagatik sortzen den errorea zuzentzeko.
 - Lagin-errorea zuzentzeko.
 - Laginaren ordezkari populazioko datuak hartzeagatik sortzen den errorea zuzentzeko.
- (12) Honako hauetatik, zeinentzako kalkulatu zenuke bereziki aniztasunaren Shannon indizea?
- Lurralde bateko lurjabetza minifundismo-latifundismo ardatzean kokatzeko.
 - ekoizpena enpresa bateko 6 lantegien artean nola banatzen den denboran zehar aztertzeko.
 - Familien errentak aztertzeko.
 - Saltokiek espazioan duten sakabanatzea neurtzeko.
- (13) Honako hauek test bateko galdera batzuen item-test korrelazioak dira: $r_{1T} = 0.2, r_{2T} = -0.4, r_{3T} = 0.3, r_{4T} = 0.8, r_{5T} = 0.99$. Cronbachen alfa 0.55 suertatu da. Galdera horietatik zein ezabatuko zenuke testaren fidagarritasuna gehitzeko?
- Lehena
 - Bigarrena
 - Bosgarrena
 - Cronbachen alfa balioa egokia da, eta ez da galderarik ezabatu behar.
- (14) Honako informazio hau daukagu etxeetako elektrogailuen gastu nominalari buruz eta elektrogailu horien prezioen bilakaerari buruz:

Urtea	2010	2011
Gastu nominala	1200	1300
Prezio indizea	140	150

Zenbat izan gastu erreala 2011an 2010eko euro konstantetan?

- 1213.3
- 1283.3
- 1352.8

(d) 1392.8

(15) Zorizko aldagai baten banaketa honela definitu da: $F(x) = x^3$; $0 < x < 1$. Kalkulatu $P[X > 0.5]$.

(a) 0.125

(b) 0.5

(c) 0.666

(d) 0.875

(16) Zorizko aldagai baten banaketa honela definitu da: $F(x) = 2x^2$; $0 < x < k$. Kalkulatu k banaketa-funtzioa izan dadin.

(a) 0.5

(b) 0.7

(c) 1.2

(d) 2.4

(17) Zorizko aldagai bat honela banatzen da:

x	0	k
$p(x)$	$1-p$	p

Kalkulatu bariantza.

(a) $k^2p(1-p)$

(b) k^2p

(c) $k^2p - kp^2$

(d) pk^2

(18) Aktibo batzuen itzaropena eta bariantza zehazten dira jarraian:

Aktiboa	A	B	C
μ	1	1.5	2
σ	3	2	1

Zein da baieztapen egokia?

(a) Epe luzera C da hobea.

(b) Epe laburrera C da hobea.

(c) Epe laburrerako aktibo egokiena zehazteko ez da behar utilitate-funtziorik.

(d) Aurreko guztiak dira zuzenak.

(19) Datuak: 2-4-5-8-10-15-20-22. Kalkulatu bigarren dezila.

(a) 2.8

(b) 3.2

(c) 3.6

(d) Aurreko guztiak faltsuak dira.

(20) Hau dakigu zorizko aldagai bati buruz: $\mu = 10, \sigma = 2$. Hurbildu $P[X < 6]$.

(a) 0.75 baino handiagoa

(b) 0.25 baino handiagoa

(c) 0.25 baino txikiagoa

(d) Aurreko guztiak faltsuak dira

Estatistika eta datuen analisisia

2019ko urtarrilaren 8a

Izena eta abizenak: Josemari Sarasola_____

Galdera	Erantzuna
1	B
2	A
3	C
4	A
5	B
6	D
7	B
8	B
9	B
10	A
11	C
12	A
13	B
14	A
15	D
16	B
17	A
18	D
19	B
20	C

KOPURUA

ONGI	10
GAIZKI	0
ERANTZUN GABE	0

ESTATISTIKA ETA DATUEN ANALISIA

Irakaslea: Josemari Sarasola

Data: 2019ko ekainaren 17an, 10:00

Iraupena: 120 min

Praktika: Adierazburu-orria

I. ebazkizuna: Zentroa eta sakabanatzea (2 puntu)

Donostia bisitatzen duten turista bikoteen artean inkesta bat egin da, besteak beste Donostian egunero egiten duten gutxi gorabeherako gastua galdetuz. Ondoren, AEBetako eta Japoniako turista bikoteen artean gastuari buruzko datuak agertzen dira, tartetan bildurik:

Gastua / Erosi al du? (\rightarrow)	Bikoteak (AEB)	Bikoteak (Japonia)
0-50	8	1
50-100	7	4
100-200	5	9
200-500	7	2
500-1000	2	2

Egin beharreko atazak:

- Gastuaren batezbesteko aritmetiko sinpleak kalkulatu bi herrialde horietako turista-bikoteetarako, eta bi herrialdeetako turista bikoteen artean, orokorrean gehien gastatzen duen taldea zein den adierazi. Zenbatekoa da diferentzia?
- Kuartil arteko ibiltartea kalkulatu AEB zein Japoniako taldeetarako.
- Aurreko neurritik abiatu, aztertu zein den gastu sabanatuena dituen turista taldea.

II. ebazkizuna: Cronbach-en alfa (2 puntu)

Test bat lau galderek osatzen dute. Jarraian, subjektu batzuek testean lorturiko puntuazio zehatzak ematen dira:

Galdera (\downarrow)/Subjektua (\rightarrow)	a	b	c	d	e	f	g	h	i	j
A	0	1	1	1	1	1	1	1	1	1
B	0	1	0	1	0	0	1	0	0	1
C	1	0	0	1	0	1	0	1	0	1
D	1	0	0	0	1	0	0	0	0	0

Egin beharreko atazak:

- Cronbach-en alfa kalkulatu.
- Lortutako emaitza interpretatu.

III. ebazkizuna: Zenbaki indizeak (2 puntu)**1. atala**

Leihoak ekoizten dituen lantegi batean zura, beira eta burdina erabiltzen dira lehengai gisa. Ondoren, azken urteotan erabilitako kopuruak eta ordaindutako prezioak jasotzen dira:

Urtea	p(zura)	q(zura)	p(beira)	q(beira)	p(burdina)	q(burdina)
2016	10	30	5	15	8	12
2017	14	42	6	10	9	14
2018	19	40	8	13	10	15

Egin beharreko atazak:

- Kalkulatu, 2016 oinarritzat harturik, 2017 eta 2018 urteetako Paascheren prezio indizeak.
- Kalkulatu 2016 oinarritzat harturik, 2017 eta 2018ko Laspeyres indize kuantikoak, eta interpreta ezazu 2017 urtekoa.

2. atala

Azken urteotan lantegian ekoizten diren leihoen sektoreko salmenta totalak eta Laspeyres prezio-indizeak jaso dira:

Urtea	2014	2015	2016	2017	2018
Salmentak	345	367	388	412	423
Indizeak (Laspeyres)	100	108	115	121	126

Salmentak deflaktatu behar dira 2015 urteari buruz. Zein urtetan saldu zen gehien termino errealetan. Deflaktatutako balioak guztiz zehatzak al dira? Zergatik?

IV. ebazkizuna: Probabilitate zuhaitza (2 puntu)

Lantegi batean eguneko 24 orduetan egiten da lana. Eguneko ekoizpena 2, 3 edo 4 unitatekoa izan daiteke 0.2, 0.5 eta 0.3ko probabilitateekin. 6 unitateko eskaera bat ekoizteko agindua jaso du.

Egin beharreko atazak:

- 6 unitate horiek (gutxienez) ekoitzi arteko probabilitate zuhaitza eratu ezazu, egunez egun.
- Adarren bukaeran adierazi lehen bi egunetako ekoizpen totala eta 6 unitate ekoizteko zenbat egun behar diren (suposatu azken honetarako ekoizpena uniformeki gertatzen dela egunean zehar).
- 2 eguneko batez besteko ekoizpena eta bariantza kalkula itzazu.
- 6 unitate ekoizteko behar den egun kopuruaren batezbestekoa eta bariantza kalkulatu itzazu.

ESTADISTIKA ETA DATUEN ANALISIA

Irakaslea: Josemari Sarasola

Data: 2019ko urtarrilaren 8a, 16:00

Iraupena: 120 min

Praktika: Ebazpena

I. EBAZKIZUNA

(a)

x	n_{AEB}	n_J	$n_{AEB}x$	n_Jx
25	8	1	200	25
75	7	4	525	300
150	5	9	750	1350
350	7	2	2450	700
750	2	2	1500	1500
	29	18	5425	3875

$$\bar{x}_{AEB} = \frac{5425}{29} = 187.06\text{€}$$

$$\bar{x}_J = \frac{3875}{18} = 215.27\text{€}$$

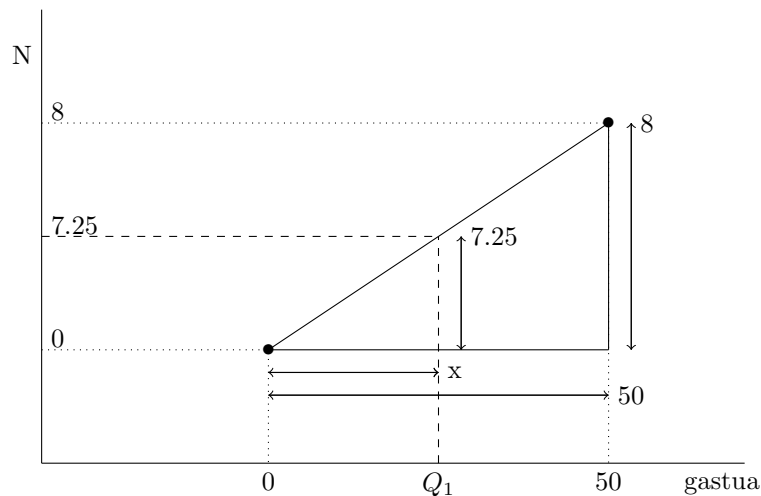
Eguneko batezbeste gehien gastatzen duen taldea japoniarrena da, eguneko 215.27€ bikote bakoitzeko, *lagin errorearen erresebapean*. Diferentzia amerikarrek 215.27-187.06=28.21€ da.

(b)

$$IQR = Q_3 - Q_1$$

AEB

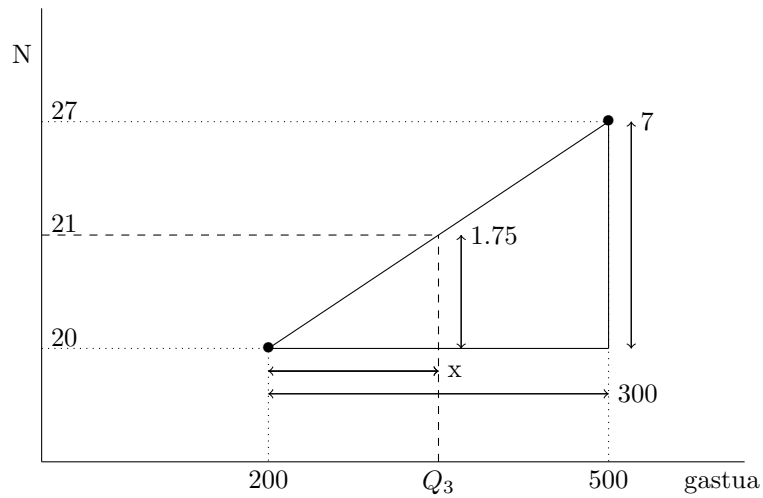
$Q_1 = 0.25 \times 29 = 7.25$ garren datua, 0-50 tartean, (8 bikote 50 arte).



$$\frac{x}{7.25} = \frac{50}{8} \rightarrow x = \frac{50}{8} \times 7.25 = 45.31 \rightarrow Q_1 = 0 + 45.31 = 45.31$$

$$IQR_{AEB} = 242.85 - 45.31 = 197.54$$

$Q_3 = 0.75 \times 29 = 21.75$ garren datua, 200-500 tartean, (8+7+5=20,8+7+5+7=27).

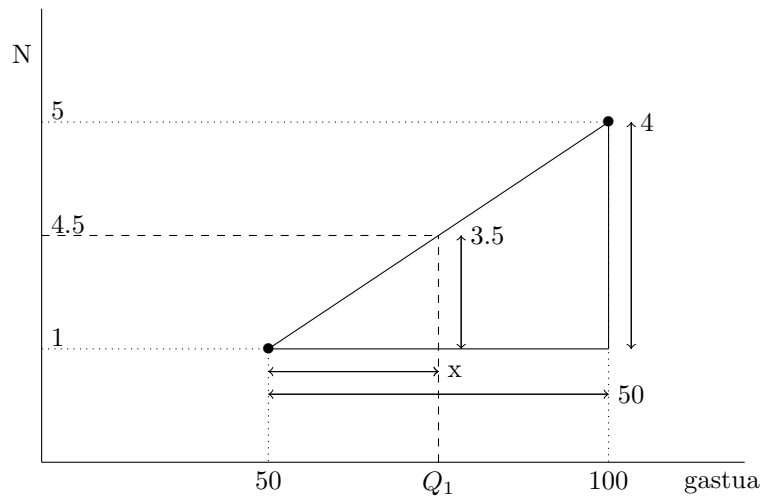


$$\frac{x}{1.75} = \frac{300}{7} \rightarrow x = \frac{300}{7} \times 1.75 = 75 \rightarrow Q_3 = 200 + 75 = 275$$

$$IQR_{AEB} = 275 - 45.31 = 229.69$$

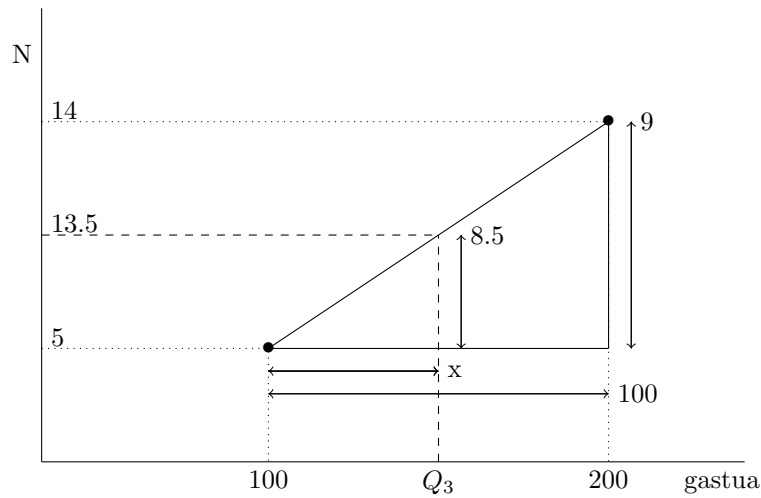
Japonia

$Q_1 = 0.25 \times 18 = 4.5$ garren datua, 50-100 tartean, (1,1+4=5).



$$\frac{x}{3.5} = \frac{50}{4} \rightarrow x = \frac{50}{4} \times 3.5 = 43.75 \rightarrow Q_1 = 50 + 43.75 = 93.75$$

$Q_3 = 0.75 \times 18 = 13.5$ garren datua, 100-200 tartean, (1+4=5, 1+4+9=14).



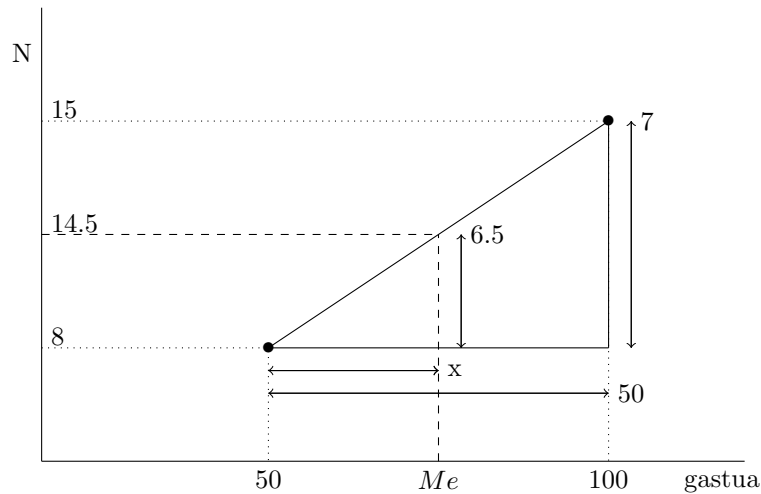
$$\frac{x}{8.5} = \frac{100}{9} \rightarrow x = \frac{100}{9} \times 8.5 = 94.44 \rightarrow Q_3 = 100 + 94.44 = 194.44$$

$$IQR_{Japonia} = 194.44 - 93.75 = 100.69$$

(c)

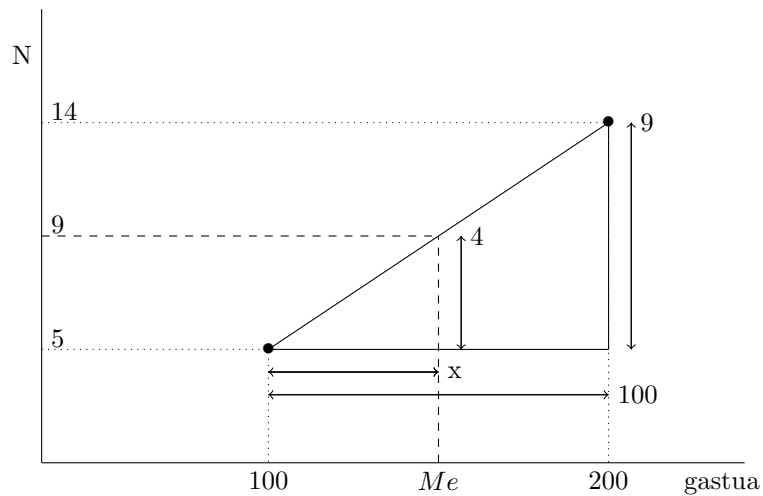
Sakabanatzea alderatzeko, neurri erlatiboak kalkulatu behar dira. IQR neurriari dagokion neurri erlatiboa IQR/Me. Medianak kalkulatu behar dira beraz:

$Me_{AEB} = 0.5 \times 29 = 14.5$ garren datua, 50-100 tartean, (8,8+7=15).



$$\frac{x}{6.5} = \frac{50}{7} \rightarrow x = \frac{50}{7} \times 6.5 = 46.42 \rightarrow Me_{AEB} = 50 + 46.42 = 96.42$$

$Me_{Japonia} = 0.5 \times 18 = 9$ garren datua, 100-200 tartean, (1+4=5, 1+4+9=14).



$$\frac{x}{4} = \frac{100}{9} \rightarrow x = \frac{100}{9} \times 4 = 44.44 \rightarrow Me_{Japonia} = 100 + 44.44 = 144.44$$

Orain, neurri erlatiboak kalkulatzeko ditugu:

AEB

$$\frac{IQR}{Me} = \frac{229.69}{93.42} = 2.45$$

Japonia

$$\frac{IQR}{Me} = \frac{100.69}{144.44} = 0.69$$

Beraz, azken emaitza horietan oinarriturik, AEBetako bikoteek gastuaren sakabanatze handiagoa dute, nabarmen gainera.

II. EBAZKIZUNA

(a)

Cronbach-en alfa kalkulatzeko item edo galdera bakoitzeko erantzunen bariantzak eta puntuazio totalaren bariantza kalkulatu behar da.

x_A	x_A^2	x_B	x_B^2	x_C	x_C^2	x_D	x_D^2	x_T	x_T^2
0	0	0	0	1	1	1	1	2	4
1	1	1	1	0	0	0	0	2	4
1	1	0	0	0	0	0	0	1	1
1	1	1	1	1	1	0	0	3	9
1	1	0	0	0	0	1	1	2	4
1	1	0	0	1	1	0	0	2	4
1	1	1	1	0	0	0	0	2	4
1	1	0	0	1	1	0	0	2	4
1	1	0	0	0	0	0	0	1	1
1	1	1	1	1	1	0	0	3	9
9	9	4	4	5	5	2	2	20	44

$$s_A^2 = \frac{9}{10} - \left(\frac{9}{10}\right)^2 = 0.09$$

$$s_B^2 = \frac{4}{10} - \left(\frac{4}{10}\right)^2 = 0.24$$

$$s_C^2 = \frac{5}{10} - \left(\frac{5}{10}\right)^2 = 0.25$$

$$s_D^2 = \frac{2}{10} - \left(\frac{2}{10}\right)^2 = 0.16$$

Eta bariantza totala:

$$s_T^2 = \frac{44}{10} - \left(\frac{20}{10}\right)^2 = 0.4$$

Eta orain Cronbach-en alfa kalkulatu dugu:

$$\alpha = \frac{4}{4-1} \left(1 - \frac{0.09 + 0.24 + 0.25 + 0.16}{0.4} \right) = -1.13$$

(b)

Negatiboa denez, ondoriozta daiteke testak ez duela inongo barne-koherentziarik. Erran nahi baita, galderek gauza ezberdinak neurtzen dituzte, ez duten neurtzen norabide berean.

III. EBAZKIZUNA

(1a)

$$P_{p,2016}^{2017} = \frac{14 \times 42 + 6 \times 10 + 9 \times 14}{10 \times 42 + 5 \times 10 + 8 \times 14} \times 100 = 132.98$$

$$P_{p,2016}^{2018} = \frac{19 \times 40 + 8 \times 13 + 10 \times 15}{10 \times 40 + 5 \times 13 + 8 \times 15} \times 100 = 173.33$$

(1b)

$$L_{q,2016}^{2017} = \frac{42 \times 10 + 10 \times 5 + 14 \times 8}{30 \times 10 + 15 \times 5 + 12 \times 8} \times 100 = 123.56$$

$$L_{q,2016}^{2018} = \frac{40 \times 10 + 13 \times 5 + 15 \times 8}{30 \times 10 + 15 \times 5 + 12 \times 8} \times 100 = 124.20$$

Hiru lehengaietatik erabilitako kopurua %23.56 emendatu da, 2017an 2016 urteari buruz, 2016 prezioetan baloraturik.

(2)

2015 urteari buruz deflaktatzeko, 2015era aldatu behar da oinarria, hirukoaren erregela sinplearen bitartez. Horren ostean, deflaktatzeko aski da kopuru nominalak oinarria aldaturiko indizeekin zatitzea:

Urtea	2014	2015	2016	2017	2018
Salmenta nominalak	345	367	388	412	423
Indizeak (oinarri: 2015)	92.59	100	106.48	112.03	116.66
Salmenta errealak	345/0.9259=372.61	367	364.38	367.75	362.59

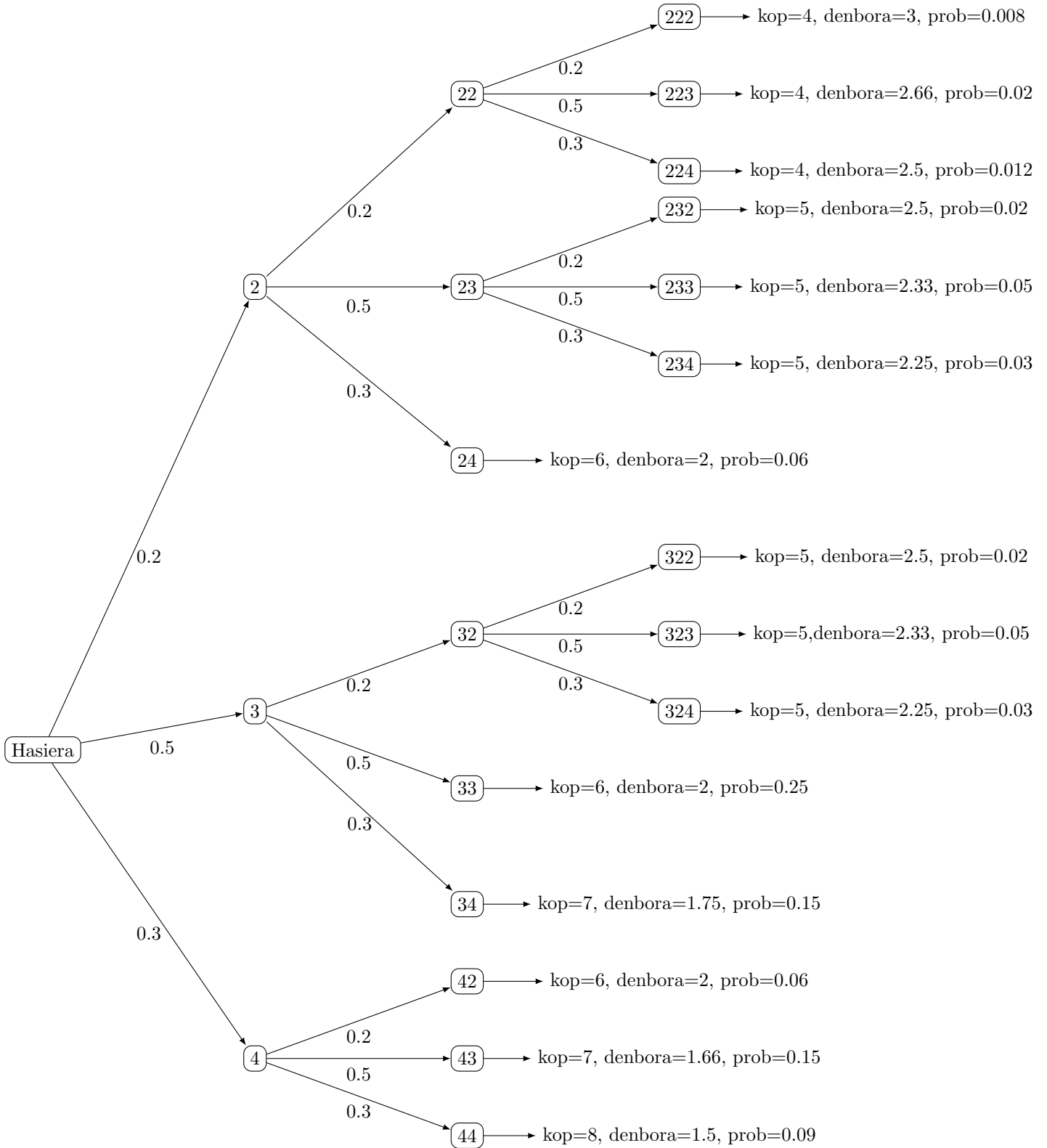
Termino errealetan 2014 urtean saldu zen gehien.

Hala eta guztiz ere, aipatu behar da deflaktaturiko balioak ez direla guztiz zehatzak, alde batetik oinarria aldatzean Laspeyres indizea zirkularra ez delako, eta bestetik Laspeyres deflaktatzaile zehatza ez delako (erabateko zehaztasuna izateko deflaktazioan Paasche indizea beharko genuke).

IV. EBAZKIZUNA

(a eta b)

Adarren hurrenez hurreneko deribazioek egunak adierazten dituzte (1., 2. eta 3. egunak). Adar bakoitzaren bukaeran 2 eguneko ekoizpena (kop), 6 unitate ekoizteko denborak eta adar-segida bakoitzeko probabilitateak (adar sinpleen probabilitateak bidertuz) zehazten dira.



(c)

x	$p(x)$	$xp(x)$	$x^2p(x)$
4	$0.008+0.02+0.012=0.04$	0.16	0.64
5	$0.02+0.05+0.03+0.02+0.05+0.03=0.2$	1	5
6	$0.06+0.25+0.06=0.37$	2.22	13.32
7	$0.15+0.15=0.30$	2.10	14.70
8	0.09	0.72	5.76
	1	6.2	39.42

$$\mu = 6.2 \text{ unitate}$$

$$\sigma^2 = 39.42 - 6.2^2 = 0.98$$

(d)

x	$p(x)$	$xp(x)$	$x^2p(x)$
1.5	0.09	0.135	0.2025
1.66	0.15	0.249	0.4133
1.75	0.15	0.2625	0.4593
2	$0.06+0.25+0.06=0.37$	0.74	1.48
2.25	$0.03+0.03=0.06$	0.135	0.3037
2.33	$0.05+0.05=0.10$	0.233	0.5428
2.5	$0.02+0.012+0.02=0.052$	0.13	0.325
2.66	0.02	0.0532	0.1415
3	0.008	0.0240	0.072
	1	1.9617	3.94

$$\mu = 1.96 \text{ egun}$$

$$\sigma^2 = 3.94 - 1.96^2 = 0.0984$$

ESTATISTIKA ETA DATUEN ANALISIA

Irakaslea: Josemari Sarasola

Data: 2019ko ekainaren 17a

Iraupena: 35 minutu

Erantzun bakarra da zuzena galdera bakoitzean. Guztira testak 2 puntu balio du. Erantzun zuzenak 0.1 puntu balio du. Erantzun oker bakoitzak horren erdia kentzen du. Galderak erantzun gabe utz daitezke, punturik gehitu eta kendu gabe.

- (1) Zertarako erabiltzen da η (eta) neurria?
- Bi aldagai kualitatiboen arteko asoziazioa neurtzeko.
 - Bi aldagai kualitatiboen arteko konkordantzia neurtzeko.
 - Aldagai kuantitatibo baten eta kualitatibo baten arteko korrelazioa neurtzeko.
 - Aldagai kuantitatiboen arteko korrelazioaren sendotasuna neurtzeko.
- (2) Zein tarteko balioak hartzen ditu kobariantzak?
- $[-1, 1]$ tartekoak.
 - Positiboak edo 0.
 - Edozein balio positibo zein negatibo.
 - Ez da inoiz 1era iristen.

- (3) Kontingentzia taula honetatik kalkulatu gamma:

Heziketa (\downarrow) / Errenta (\rightarrow)	Txikia	Handia
Txikia	4	8
Handia	6	3

- 0.6
 - 0.6
 - 0.4
 - 0.4
- (4) Errenta datuak: 2-2-2-2-3-3-3-5-5-8-8. Kalkulatu Robin Hood adierazlea (portzentajejan).
- 25
 - 40
 - 50
 - 60
- (5) Errenta datuak: 2-2-3-5-8. Kalkulatu mediala.
- 3
 - 4
 - 5
 - 8
- (6) Lorenz kurbako puntuak dira hauek: $(p_1 = 0.1, q_1 = 0.01)$; $(p_2 = 0.4, q_2 = 0.3)$; $(p_3 = 0.6, q_3 = 0.45)$; $(p_4 = 0.9, q_4 = 0.7)$. Kalkulatu Palma ratioa.
- 2.33
 - 2
 - 1.33
 - 1
- (7) Enpresa batek 10 denda frankiziatu ditu. Denda horietako fakturazioak irudikatzeko zein da diagrama egokia?
- Puntu-diagrama
 - Kaxa-diagrama
 - Histograma

- (d) (a) eta (b) zuzenak dira.
- (8) 30 datuetatik zera dakigu: $IQR = 8$; $Q_3 = 16$. 4 datu handienak honako dira: 20-21-26-27. Zein da egia kaxa-diagramari buruz?
- (a) Albo-balioa 21 da eta muturreko datuak 26 eta 27.
 - (b) Albo-balioa 24 da eta muturreko datuak 26 eta 27.
 - (c) Albo-balioa 27 da eta ez dago muturreko daturik.
 - (d) Albo-balioa 28 da eta ez dago muturreko daturik.
- (9) 10 datuetatik populazio-bariantza 5 da. Kalkulatu lagin-bariantza.
- (a) 5.55
 - (b) 2.35
 - (c) 4.5
 - (d) 2.12
- (10) Zertarako erabiltzen da korrelazio partziala?
- (a) Aldagai biren arteko korrelazioan, muturreko datuen eragina baztertzeke.
 - (b) Aldagai biren arteko korrelazioan, hirugarren aldagai baten efektua baztertzeke.
 - (c) Korrelazioa kalkulatzeko, baina soilik aldagai biak dikotomiko direnean.
 - (d) Sasikorrelazioa (korrelazio faltsua edo fiktizioa) baztertzeke.
- (11) Zein da banaketa-funtzio batek bete behar EZ duen baldintza?
- (a) $F(sup) = 1$
 - (b) $F(inf) = 0$
 - (c) Azpiko azalera 1 izan behar da.
 - (d) Gorakorra izan behar da.
- (12) $F(x) = \frac{x^2}{4}; 0 \leq x \leq 2$. Zenbat da $P[X < 1]$?
- (a) 0.125
 - (b) 0.25
 - (c) 0.5
 - (d) 0.75
- (13) $F(x) = \frac{x^2}{4}; 0 \leq x \leq 2$. Zenbat da $P[0.5 < X < 1.5]$?
- (a) 0.125
 - (b) 0.25
 - (c) 0.5
 - (d) 0.75

- (14) $f(x) = \frac{7x^2}{90}; 0 \leq x \leq 3$. Zenbat da $P[X < 1]$?
- (a) 0.02 gutxi gorabehera
 - (b) 0.12 gutxi gorabehera
 - (c) 0.22 gutxi gorabehera
 - (d) 0.32 gutxi gorabehera
- (15) $f(x) = \frac{7x^2}{90}; 0 \leq x \leq 3$. Zenbat da itxaropen matematikoa?
- (a) 1.57
 - (b) 1.67
 - (c) 1.77
 - (d) 1.87
- (16) A eta B inbertsioek eguneko ematen duten mozkinen buruz hauek dakizkigu: $\mu_A = 4, \sigma_A = 1; \mu_B = 6, \sigma_B = 2$. Zein da egiazkoa honako hauetatik?
- (a) Epe luzera hobe da A, eta B arriskutsuago da gainera.
 - (b) Epe luzera hobe da B, baina B arriskutsuago da.
 - (c) Epe luzera hobe da A, baina A arriskutsuago da.
 - (d) Epe luzera hobe da B, eta A arriskutsuago da gainera.
- (17) Nola deitzen zaio XIX. mendera arte besterik gabe datu ekonomiko eta administratiboak biltzen zituen eskola estatistikoari?
- (a) Aritmetika politikoa
 - (b) Estatistika politikoa
 - (c) Esplorazio estatistikoaren garaia
 - (d) Zentsura estatistiko-politikoa
- (18) Nola deitzen zaio psikologia arloan bereziki garatzen den estatistikaren arloari?
- (a) Psikoestatistika
 - (b) Psikometria
 - (c) Datuen psikoanalisi
 - (d) Psikoteknika
- (19) Honako neurri estatistiko hauetatik zein da sendoa edo jasankorra?
- (a) Ibiltarrea
 - (b) Desbideratze estandarra
 - (c) IQR
 - (d) Aldakortasun koefizientea
- (20) Banaketa diskretu bati buruz: $P[X = x] = \frac{x}{10}; x = 1, 2, 3, 4$. Zenbat dira $P[X \leq 3]$ eta itxaropena?
- (a) 0.3 eta 3, hurrenik hurren.
 - (b) 0.3 eta 2, hurrenik hurren.
 - (c) 0.6 eta 3, hurrenik hurren.
 - (d) 0.6 eta 2, hurrenik hurren.

Estatistika eta datuen analisisia

-
2019ko ekainaren 17a**Izena eta abizenak: Josemari Sarasola**_____

Galdera	Erantzuna
1	c
2	c
3	b
4	a
5	c
6	d
7	d
8	c
9	a
10	b
11	c
12	b
13	c
14	a
15	a
16	b
17	a
18	b
19	c
20	c

KOPURUA

ONGI	10
GAIZKI	0
ERANTZUN GABE	0