

ESTADÍSTIKA 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 neurri 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 modeloa (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 likidatuko 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 errorearen eta antzeko ikerketen erreserbapean betiere.

II. EBAZKIZUNA

(a): Khi-karratu

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

87	92	179
157	98	255
244	190	434

Maiztasun teorikoak (E) kalkulatzeko ditugu:

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

 $(O - E)^2/E$ kalkulatzeko 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 kalkulatzeko 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 nabarmenena dela esango genuke, $(O - E)/\sqrt{E}$ balioei errepatuz (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 dugun:

4 ordenagailu erositako,

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

5 ordenagailu erositako,

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 erositako,

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 erositako,

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 erositako,

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 erositako gertatzen da. Beraz, 7 ordenagailu erosi behar dira.

(b)

7 ordenagailu erositako,

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

Beraz, batezbeste 0.15 ordenagailuren salmenta galduko da.

(c)

7 ordenagailu erositako,

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 estandarraren 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)?
 - (a) 1
 - (b) 2

- (c) 3
(d) 4
- (8) Zenbat da y aldagaiaren kuasibariantza edo bariantza zuzendua?
(a) 1
(b) 2
(c) 3
(d) 4
- (9) Zenbat da x eta y aldagaien arteko kobariantza?
(a) -1
(b) -2
(c) 1
(d) 2
- (10) Pearson korrelazio koefiziente lineala kalkulatu, nola interpretatu behar da?
(a) Korrelazioa alderantzizkoa eta perfektua da.
(b) Korrelazioa alderantzizkoa eta sendoa da.
(c) Korrelazioa zuzena eta perfektua da.
(d) Korrelazioa zuzena eta sendoa da.
- (11) Zertarako erabiltzen da kuasibariantza edo bariantza zuzendua?
(a) Bariantza datu atipiko edo muturrekoen eraginetik aparte mantentzeko.
(b) Datuak tartetan izanik, erdipuntua hartzeagatik sortzen den errorea zuzentzeko.
(c) Lagin-errorea zuzentzeko.
(d) Laginaren ordeztasuneko datuak hartzeagatik sortzen den errorea zuzentzeko.
- (12) Honako hauetatik, zeinentzako kalkulatu zenuke bereziki aniztasunaren Shannon indizea?
(a) Lurralde bateko lurjabetza minifundismo-latifundismo ardatzean kokatzeko.
(b) Ekoizpena enpresa bateko 6 lantegien artean nola banatzen den denboran zehar aztertzeko.
(c) Familien errentak aztertzeko.
(d) 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?
(a) Lehena
(b) Bigarrena
(c) Bosgarrena
(d) 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?
(a) 1213.3
(b) 1283.3
(c) 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