

# DATU ANALISIA ETA SAIKUNTZA DISEINUAK

## PORTAERA ZIENTZIETAN

Xabier Isasi, Nekane Balluerka  
Udako Euskal Unibertsitatea, Bilbo, 1995

Interes gune den fenomeno eza-gutzeaz gain, Portaera Zientzien eremuan ikerkuntza egin nahi duen profesional orok, baliabide metodologikoak behar bezala erabili behar ditu bere lana modu zorrotzean egin nahi badu. Hemen aurkezten den liburua ikerketaren azken faseari, datuen analisiari alegia, lotzen zaio batik bat. Bertan, egileek, estatistika inferentzialari dagozkion oinarrizko gaiak jorrazten dituzte eta, bide batez, Portaera Zientzien eremuan gehienbat erabiltzen diren saiakuntza diseinuen aurkezpena egin eta berauei lotutako bariantza analisiak prozedura desberdinen bitartez garatzen dituzte. Unibertsitate mailako ikasleei zuzendutako oinarrizko testuliburua izanik, adibide anitzez hornitzen da gai bakoitza; honela, ulergarritasunari eta ikuspegi didaktikoari garrantzi berezia eman zaiola azpimarratu nahi nuke. Bestalde, liburu osoan zehar, Portaera Zientziei egokitutako estatistikaren eta saiakuntza diseinuen eremuei dagoen hizkuntza teknikoak finkatzeko eta egokitzeko egileek egin duten ahalegina ere gutziz aipagarria litzateke.

Liburuan bi zati nagusi bereiz daitezke. Lehen zatian (lehenengo bederatzi kapituluak) estatistika inferentzialari dagozkion oinarrizko gaiak aztertzen dira; hots, probabilitatea eta zorizko jolasak, matrizeak eta determinanteak, zorizko aldagaien

banaketak, laginketa, estimazioa, hipotesi testak, test ez-parametrikokoak, lerro eredu orokorra eta bariantza analisia. Bigarren zatian, berriz, konplexutasun txikiagotik handiagorako irizpideari jarraituz, Portaera Zientzien eremuan gehienbat erabiltzen diren saiakuntza diseinuak azaltzen zaizkigu eta bakoitzari dagokion bariantza analisia metodo desberdinen bidez aurkezten zaigu.

Oraintxe aipatutako egiturari helduz, egileak probabilitateaz hitz egiten hasten dira. Lehen kapitulu honetan, esperimenduak eta zorizko gertaerak, lagin espazioa eta probabilitatea bera definitzen dizkigute. Halaber, probabilitatearen axiomak zeintzuk diren aipatu eta zenbatze teknika nagusiak aurkezten dituzte. Atal honetan bertan, probabilitateen bi lege nagusiak, batera ezintasunarena eta independentzia-rena hain zuzen, eta oinarrizko teorema, probabilitate batuketarena, probabilitate biderkaketarena eta probabilitate osoarena alegia, aztertzen dira. Probabilitatearen esparruan hain garrantzitsu diren itxaropen matematikoen kontzeptua, zenbaki handien legea, baldintzazko probabilitatea eta zorizko aldagaien kontzeptua azalduz amaitzen da atala.

Bigarren kapituluan matrizeak eta determinanteak aztertzen dira. Matrizea zer den agertu ondoren, egileek matrize mota nagusiak eta matrizee-

kin egin daitezkeen eragiketak aurkezten dizkigute. Halaber, atal honetan determinantearen kontzeptua, bere propietateak eta berorren kalkulua burutzeko erabide desberdinak aurkitzen ditugu. Matrize eta determinanteetarako hurbilpen orokorra eskaintzeaz gainera, egileek gai hau berau datu analisiaren testuinguruan kokatzen dute.

Matrize eta determinantei buruzko atalaren ondoren, zorizko banaketak azaltzen dizkigute. Hirugarren kapitulu honetan, zorizko aldagai diskretuen nahiz jarraien banaketak zertan dautzan aurkitzen dugu. Zorizko aldagai diskretuei dagokienez, banaketa uniforme, binomiala eta multinomialaren kontzeptua eta probabilitate funtzioak azaltzen dizkigute. Zorizko aldagai jarraien kasuan, berriz, banaketa normala, Pearsonen  $\chi^2$  banaketa, F motako banaketak eta t motako banaketak aldagai diskretuei dagozkienak baino xehekiago aztertzeo aukera ematen digute.

Laugarren atalean laginketa lantzen da funtsean. Laginketa ukitu aurretik, Isasik eta Balluerkak informazio arazoaren eta ikerkuntza arazoaren arteko bereizketa planteatzen dute eta, bereizketa honen ildotik saiakuntza diseinuetarako hurbilpen orokorra egiten dute. Jarraian, laginketa garatzen dute. Laginketaren barnean hain garrantzitsua den zorizkotasunaz mintzatzeko lehenik eta, ondoren, Portaera Zientzietan zorizko laginketarako erabiltzen diren prozedura nagusiak azaltzen dizkigute. Populazioaren batazbestekoa estimatu ahal izateko beharrezko laginaren tamainaren kalkulua aurkeztuz amaitzen da atala.

Edozein estimazioren abiaburua eta oinarria den laginketari aurreko atala eskaini ondoren, egileak estimazioari berari hurbiltzen zaizkio. Honela, zenbaketaaren auzia eta errore estatistikokoaren kontzeptua aurkezten dituzte lehenik eta, bigarrenik, estimagailuen funtsezko ezaugarriak ezagutarazten. Kapitulu honetan bertan estimagailuen laginketa banaketa, populazio batazbestekoaren estimagailua eta batazbesteko estimagailuaren bariantza zertan dautzan azaltzen dute.

Seigarren kapitulua hipotesi testei dagokie. Kapituluaren mamira iritsi aurretik, irakurleak hipotesi orokorren, hipotesi zientifikoen eta hipotesi estatistikoen arteko desberdintasunak aurki ditzake. Jarraian, hipotesi estatistikotatik bete-betean murgil daiteke. Lehenik, hipotesi estatistiko bakunen eta konposatuaren arteko bereizketa azter dezake eta, ondoren, hipotesi nulua eta ordezkio hipotesiaren artekoa. Halaber, hipotesi testaren kontzeptua, berau burutzeko prozedura orokorra, test estatistiko zuzena hautatzeko irizpideak eta hipotesi testei atxekiriko erroreak, I. motako errorea eta II. motako errorea, alegia, zertan dautzan ikas dezake. Jarraian testaren indar kontzeptua aurki dezake eta, azkenik, zentru joerako estatistikoa erabiliaren, hots, batazbesteko aritmetikoaren inguruko hipotesi testak eta bariantza eta proportzioari dagozkion testak ere azter ditzake.

Aurreko ataleko edukinak azaldu ostean, egileek test ez-parametrikotara hurbiltzen gaituzte. Lehenik, test parametrikoen eta ez-parametrikoen erabilerarako aurretikoak eta irizpideak ematen dizkigute eta, ondoren, test ez-parametrikoko ezagunenen helburua eta beren erabilerarako baldintzak labur-

biltzen dituzte. Modu berean, test ez-parametrikoen esparruan garrantzi berezia duen kontingentzia kontzeptua azaltzen dute eta test ez-parametrikotzat batzuen garapena, hala nola lagin bakarreko Chi karratu testarena eta Kolmogorov-Smirnov delakoarena. Atal honetan bertan, Isasik eta Balluerkak kapitulu honen eta aurreko kapituluaren (6. atalaren) osagarritzat har daitekeen taula bat aurkezten digute. Taula hau guztiz erabilgarri gerta dakiguke, zeren berorretan estatistika inferentzialeko parametro nagusien inguruko hipotesi testak burutzeko parametroak eta hipotesiak, lagina(k), testen oinarriak eta dagozkien banaketak eta estatistikoak eskaintzen baitizkigute.

Zortzigarren kapituluan, Giza eta Gizarte Zientzietako ikerkuntzan maizen erabiltzen diren analisisen oinarrian dagoen lerro eredu orokorraz luze mintzatzen zaizkigu. Bertan, lerro ereduari dagozkion analisi teknika ezagunenetariko bi azaltzen dizkigute: lerro erregresio bakuna eta lerro erregresio anizkoitza, hain zuzen ere. Bakoitzaren azalpen sakona egiteaz gainera, lerro ereduari dagokion edozein analisi erabiltzeko betebeharreko aurretikoak aurkezten dizkigute.

Bederatzigarren atalean, bariantza analisisirako hurbilpen orokorra aurkitzen dugu. Bertan, egileek, bariantzen deskonposaketa oinarritzat hartuz, bariantza analisisia zertan datzan azaltzen digute. Hurrengo atalerako sarrera gisa erabilgarria gerta dadin, kapitulu honetan berreduren baturak, libertate graduak, berreduren batzbestekoak eta F estatistikoaren kontzeptuak definitzen dituzte.

Hamargarren kapituluak berak liburuaren bigarren zatia eratuko luke.

Izan ere, 150 orrialde inguruko atala dugu. Atal honetan, Balluerkak eta Isasik Portaera Zientzien eremuan sarrienik erabiltzen diren diseinuak azaltzen dizkigute, hots, faktore bakarreko taldearteko diseinua, taldearteko diseinu faktorialak, taldebarneko diseinu esperimental bakuna, bi faktorez osotutako taldebarneko diseinu faktoriala, diseinu faktorial mistoa, aldagai kobariantedun diseinua eta ausazko blokeko diseinua. Diseinu bakoitzaren azalpenarekin batera, bere abantailak eta desabantailak, berari dagokion bariantza analisiaren egitura eredu matematikoa eta bariantza analisiaren kalkulua eskaintzen digute. Azken puntu honi dagokionez, guztiz azpimarragarria da bariantza analisiaren berreduren baturak ohizko prozeduraz garatzeaz gainera, bektoreen bidezko ebazpena ere aurkezten dutela. Prozedura hau, kalkulurako ohizko algoritmoak baztertuz, diseinu bakoitzari dagokion egitura eredu matematikoaren parametroetatik abiatzen da eta, ondorioz, analisisia askoz didaktikoagoa eta ulergarriago gertatzen da.

Amaitzeko, egileek, datu analisisirako gida praktikoa bat (XI. kapitulua) eta zenbait eraskin gehitzen dizkiote testuari. Lehen eraskinean, irakurleak, liburuan aztertutako diseinu bakoitzari dagokion bariantza analisisia ordenagailuz egin ahal izateko SPSS-eko programak eta datu analisiaren irteerak aurki ditzake. Bigarren eraskinean, matematika eta estatistikaren eremuan gehien erabiltzen diren letra grekoak azaltzen dizkigute. Jarraian, Portaera Zientzietan estatistikarako taularik erabilgarrienak ikus daitezke eta, azkenik, datu analisiaren eta saiakuntza diseinuen eremuetan gehien erabilitako kontzeptu eta hitz teknikoak zerrenda dakarkigute.

Laburbilduz, liburua Portaera Zientziei egokitutako estatistikari eta saiakuntza diseinuei dagozkien oinarritzko gaiak aztertzeko guztiz erabilgarria dela deritzot. Gainera, ikuspegi didaktikoari emandako garrantzia kontutan hartuz, lizentziatura nahiz hirugarren zikloko ikasle euskaldunentzat oso

egoki gerta daiteke. Azken batez, metodologian interesaturik dagoen edo metodologiaren garrantzia onartzen duen orok irakurri beharko lukeen lana dugu.

*Juan Etxeberria Murgiondo*