

Estatistika zer den

- Zer? (objektu materiala) Fenomeno aldakorrei buruzko datuak.
- Nola? (objektu formal) Haietan dauden erregulartasunak bilatuz.
- Konklusioak beti orokorrak, ez indibidualak.
- Ez da datu bilketa hutsa, zer egin nahi den garbi eduki behar da.
- Estatistika zientziaren paradigma positibistari dagokio.
- Estatistika ikerketa nomotetikoak garatzeko erabiltzen da.
- Indukzio izeneko metodo zientifikoarekin loturik dago: kasu partikularrak → lege orokorrak.

Estatistikaren historia

- Antzinetatik zentsuak eta bestelako datu bilketak egin dira.
- Lehen azterketa estatistikoa: John Graunt kapitaina, XVII. mendean, Londresko jaiotza eta heriotzei buruz.
- Hasiera batean, datu geografiko eta ekonomikoak biltzen dira: aritmetika politikoa.
- XVIII. mendea: Achenwall-ek *Statistik* hitza asmatzen du.
- XVII. mendean, probabilitate kalkulua abiarazten da.
- XIX. mendea, probabilitatea eta datuen azterketa uztartu eta estatistika modernoa sortzen da.
- XIX. mendea: giza eta gizarte zientzietara zabaltzen da (Adolphe Queteletak batez besteko gizakiaren ideia zabaltzen du).
- Estatistikaren zirkulua XX. mendean ixten da, metodo estatistikoaren oinarri teorikoak finkatuz.
- XX-XXI. mendea: informatikaren garapenarekin batera, *big data* eta *data mining* arloak garatzen dira.

Errore estatistikoa. Estatistika deskribatzailea eta inferentzia estatistikoa

Eraitza estatistikoetan, adibidez batezbesteko bat kalkulaturik, bi motako erroreak izaten dira:

- ikertu nahi den populazio osoaren ordez, horren adierazgarri gisa hartzen den zorizko lagin bateko datuak hartzen direnean, laginketa-errorea deitzen dena;
- jasotako datuak zorizkoak direla (berez edo faktore kontrolagaitzak izateagatik) kontuan hartzen duena, zorizko errorea edo errore estokastikoa deitzen dena (dadoa jaurtirik, ez dakigu zein zenbaki izango den).

Errore estatistikoa. Estatistika deskribatzailea eta inferentzia estatistikoa

- Datu-multzo bat bat errorea kontuan hartu gabe aztertu egiten denean, **BESTERIK GABE, ESTADISTIKA DESKRIBATZAILEA** egiten dela esaten da.
- Azterketan errorea kontuan hartzen denean, **INFERENTZIA ESTADISTIKOA** dela esaten da.

Azterketa estatistiko bat zorrotz egin nahi bada, inferentziako teknikak erabili behar dira. Baina horiek konplexuagoak dira, probabilitate-teoria eta baliatzen dutelako. Azterketa sinple eta esploratzaile bat bilatzen bada, estatistika deskribatzailearekin nahikoa da.

Estatistika zientzia eta teknikaren arlo guztietan erabiltzen da. Arlo jakin batean teknika estatistiko bereziak erabiltzen dira maiz, eta horrela arlo horretako estatistikari izen berezi bat ematen zaio:

- ekonometria, ekonomiari aplikaturiko estatistika;
- psikometria, psikologiari aplikaturiko estatistika;
- bioestatistika, biologiari eta medikuntzari aplikaturiko estatistika;
- epidemiologia, gaixotasunen maiztasuna aztertzen duen estatistika;
- geoestatistika, geologiara aplikaturiko estatistika.
- demografia, astrometria, ...

Zer jakin nahi dugun zehaztu:

- Konparaketak egin datu multzo ezberdinen artean, edota
- aldagaien arteko erlazioak bilatu, edota
- auresanak egin.

Halaber, populazio osoa aztertu (zentsu bat, alegia) edo horren azpimultzo bat, adierazgarri izan dadin zoriz jaso beharko dena, jaso eta aztertu den (lagin bat aukeratu eta aztertu, alegia).

Datuak aztertu aurretik, horiek modu egokian jasotzeko lana planifikatu behar da:

- ikerketa-unitatea (zer objektu ikertuko den) eta behaketa-unitatea (datua zeren gainean jasoko den) zehaztu behar dira (adibidez, familien errenta aztertu nahi bada, ikerketa-unitatea familia izango da eta behaketa-unitatea pertsona fisikoa (familiako kideen errenta jaso behar delako familiako errenta jakiteko), eta zehatz-mehatz definitu beharko da zer den familia bat, datu guztiak homogeneoak izan daitezzen);
- espazio-eremua (nongo errentak?);
- denbora-eremua (noizko errentak);

Halaber, populazio osoa aztertu (zentsu bat, alegia) edo horren azpimultzo bat, adierazgarri izan dadin zoriz jaso beharko dena, jaso eta aztertu den (lagin bat aukeratu eta aztertu, alegia).

Aldagaia datuaren bitartez jaso edo neurtu egiten dena da (adibidez, errenta, adina).

- neurtu gabe, ezaugarri bat *adierazi* egiten denean, aldagaia KUALITATIBOA edo KATEGORIKOA da (aukeraturiko gradua).
- *neurtu* egiten denean, aldagaia KUANTITATIBOA da (altuera zmtan); zenbat? galderari erantzuten diote.
- mailakatu egiten denean, aldagaia ORDINALA dela esaten da batzuetan (errenta altua/ertaina/txikia).

Aldagaiak ez dira berez kuantitatibo edo kualitatibo, nola jasotzen diren adieraziko digu nolakoak diren. Adibidez, matematika-nota kuantitatiboa (8.7) eta kualitatiboa (oso ongi) izan daiteke.