

Fisher proba zehatza

Josemari Sarasola

Estatistika enpresara aplikatua

Gizapedia



Banaketa hipergeometrikoan oinarriturik, **kontingentzia taula** batean bi aldagaien artean independentzia dagoen probatzeko erabiltzen da. Adibidez, kontingentzia taula honetan gaingituen eta ez gaingituen proportzioak emakume eta gizonen artean berdinak (edo desberdinak) diren probatzeko erabiltzen da:

Sexua ↓ / Emaizta →	Gaingitua	Ez gaingitua	Guztira
Gizona	8	5	13
Emakumea	3	12	15
Guztira	11	17	28

Hipotesi nuluan bi aldagaiak independenteak direla ezartzen da, hau da, gizon eta emakumeengan ez dagoela proportzio diferentziarik gaitutuen eta ez gaitutuen artean.

Independentziaren kasuan egongo liratekeen maiztasun teorikoak hauek lirateke, gelaska bakoitzean bazter maiztasunak bidertuz eta lagin tamainarekin zatituz (adibidez, $(11 \times 13)/28 = 5.1$):

Sexua ↓ / Emaizta →	Gaitutua	Ez gaitutua	Guztira
Gizona	5.1	7.9	13
Emakumea	5.9	9.1	15
Guztira	11	17	28

Maiztasun teorikoak eta enpirikoak alderatuz, gizonek maizago gairitzen dutela dirudi, eta beraz hori da hipotesi nuluari dagokion aukerako hipotesia. Harritzeko modukoa, beraz, goitik dago gelaska horretan: gairitu duten 8, 9, 10 edo 11 gizon. Horren probabilitatea kalkulatu behar da.

Hipotesi nulupean independentzia dago, eta beraz 28 pertsonetatik aukeratutako 11 aprobatuak zoriz banatzen dira gizon eta emakumeen artean, jakinda betiere 13 gizon eta 15 emakume daudela; edota, beste norabidean, 20 pertsonetatik aukeratutako 13 gizonak zoriz banatzen dira gairituen eta ez gairituen artean.

Hori horrela izanik, gainditu duten 8, 9, 10 edo 11 gizon izateko probabilitatea honela kalkulatu daiteke:

- $$P[8, 9, 10, 11] = \frac{\binom{11}{8}\binom{17}{5}}{\binom{28}{13}} + \frac{\binom{11}{9}\binom{17}{4}}{\binom{28}{13}} + \frac{\binom{11}{10}\binom{17}{3}}{\binom{28}{13}} + \frac{\binom{11}{11}\binom{17}{2}}{\binom{28}{13}} = 0.0309$$

- $$P[8, 9, 10, 11] = \frac{\binom{13}{8}\binom{15}{3}}{\binom{28}{11}} + \frac{\binom{13}{9}\binom{15}{2}}{\binom{28}{11}} + \frac{\binom{13}{10}\binom{15}{1}}{\binom{28}{11}} + \frac{\binom{13}{11}\binom{15}{0}}{\binom{28}{11}} = 0.0309$$

Adierazgarritasun-maila %5 hartzen bada, hipotesia nulua baztertu behar da, p-balioa (0.0309) txikiagoa delako, eta sexua eta azterketa gainditzea loturik daudela ondorioztatu.

Ohartu behar da proba alde bakarrekoa dela kasu honetan, aukeran gizonek maizago gainditzen dutela onartzen dutelako (eta ez soilik sexuak gainditzeari eragiten diola). Eta horregatik alderatu behar da p balioa α -rekin, eta ez $\alpha/2$ balioarekin.

Zalantza sor daiteke Fisherren proba zehatza ezkerreko goiko izkinako maiztasuna oinarritzat harturik, edota beste gelaska bateko maiztasuna harturik ere gara daitekeen.

Arrazoirik gabeko zalantza da: taulako zutabeak eta errenkadak elkar aldatu egin daitezke, eta beraz 4 gelasketatik edozein har daiteke oinarritzat kalkuluak egiteko: banaketa hipergeoemtrikoaren 4 simetriengatik, emaitzak beti berdinak izango dira.

Kalkuluak laburtzeko halere, hobe da maiztasun txikiena hartzea.

Zergatik deitzen zaio Fisher proba *zehatza*? Independentzia probatzeko beste test estatistikoen aldean, erabateko zehaztasunez kalkulatu duen p -balioa. Aukeran dauden beste probak (khi-karratua oinarritzen dena, esaterako) hurbilketak dira, maiztasunak berez diskretuak izanda, jarraituak direla suposatzen dutelako. Maiztasunak beti diskretuak dira, baina bereziki diskretuak maiztasunak txikiak direnean. Eta horregatik, Fisherren proba zehatza bereziki maiztasun txikien kasuan erabiltzen da.

Fisher probak eragozpenak baditu, ordea:

- Finkotzat hartzen ditu bazter-maiztasunak, suposizio nahiko itxia dena (zergatik hartzen dira aldakortzat barrukoak, eta finkotzat bazterrekoak?)
- Proba kontserbatzailea da, hipotesi nulua, independentzia alegia, onartzeko joera duena, bereziki maiztasun txikietarako, haietan zaila baita α baino probabilitate txikiagoak izatea. Horregatik, komeni da proba honetan ohi baino adierazgarritasun-maila handiagoak ezartzea aurretik, hipotesi nulua aiseago baztertzeko.