Revista de Dirección y Administración de Empresas. Número 10, diciembre 2002 págs. 147-149
Enpresen Zuzendaritza eta Administraziorako Aldizkaria. 10. zenbakia, 2002 abendua 147-149 orr.

ZORIZKO FLUXU BATI EZARRITAKO MUGAREN ZENBATESPENA EGIANTZ HANDIENAZ

JOSEMARI SARASOLA LEDESMA
Donostiako Enpresa Ikasketen Unibertsitate Eskolako irakaslea

1. ADIBIDE BATEN EBAZPENA

Enpresa bateko tekniko informatikoak bi motako lanak egiten ditu enpresa batean: langileek beren ordenagailuetan dituzten arazoak konpondu eta enpresako web orria kudeatu. Langileek planteatzen dituzten arazoen kopurua Poisson-en banakuntzaren araberakoa da, batezbestekoa 4 izanik. Hala ere, teknikoak k muga bat ezartzen du konpontzen dituen arazoetarako, web orriak gutxieneko denbora bat behar duelako. Argibidez, k=4 bada, 6 arazoetarako laguntza eskatzen bazaio, 4 arazo bakarrik atendituko ditu. k muga parametro ezezaguna da. Hainbat egunetan zehar teknikoak konpondu egin dituen arazoen kopurua jaso da:

3 2 3 4

k parametroa zenbatetsi behar da egiantz handienaren ihardunpidea erabiliz.

Teknikoak konpondu egiten dituen arazo kopuruari buruzko datuei dagokien egiantz funtzioa eratu behar da. Horretarako garbi utzi behar da aldez aurretik zenbatetsi nahi dugun k muga 4 edo handiagoa izan behar dela dela, bestela ezingo litzateke 4 balioa hartzen duen datua gertatu; beste era batera esanda, k parametroaren espazio parametrikoa $k \geq 4$ da.

Egiantz funtzioa eman dezagun orain. Jakina denez, egiantz funtzioak mostra jakin bat gertatzeko probabilitatea ematen du parametroaren balio ezberdinetarako. Izenda ditzagun ‘Y: teknikoak konpondu egiten duen arazo kopurua’ eta ‘X: teknikoari planteatutako arazo kopurua’. Jasotzen dugun mostra Y aldagaiari buruzkoa da.

$$L(Y;k)=P[Y=3]\cdot P[Y=2]\cdot P[Y=3]\cdot P[Y=4]$$

Teknikoak 0, 1, 2 edo 3 arazo konponduko ditu 0, 1, 2 edo 3 arazo planteatzen zaionean, k muga gutxienez 4 baita. Beraz,

$$L(Y;k)=P[Y=3]\cdot P[Y=2]\cdot P[Y=3]\cdot P[Y=4]=P[X=3]\cdot P[X=2]\cdot P[X=3]\cdot P[Y=4]$$

X aldagaiari buruzko probabilitateak guztiz zehaztuta daude, P(λ=4) prozesu bati dagozkionak baitira, eta ez daude k parametroaren menpe. Ondorioz, egiantz funtzioaren balioa P[Y=4] probabilitatearen balioaren araberakoa izango da. Beraz, egiantz funtzioa maximotzen duen k balioa aurkitzeko, P[Y=4] probabilitatea maximotzen duen k balioa aurkitu behar dugu:

$$P[Y=4]=\begin{cases} P[X=4], & k>4 \\ P[X\geq 4], & k=4 \end{cases}$$

Izan ere, 4 arazo <u>konpondu</u> egingo ditu, k muga 4 baino handiagoa izanik, 4 arazo planteatu egiten zaionean. Aitzitik, k muga 4 izanik, 4 arazo konpondu egingo ditu, 4 arazo edo gehiago <u>planteatu</u> egiten zaionean.

Argi dago $P[X\geq 4]>P[X=4]$ betetzen dela. Beraz, $P[Y=4]$, eta bide batez egiantz funtzioa, maximotu egiten duen k balioa k=4 da. Beraz, hauxe izango da egiantz handieneko zenbatespena kasu honetarako: $\hat{k}_{EH}=4$

2. **EBAZPEN OROKORRA**

Eman dezagun X aldagai bat, sinplifikatzeko, eta orokortasuna galdu gabe, diskretu pentsatuko duguna. X aldagai honen gainean aldagai aldaketa hau egiten badugu, Y aldagai berria sortzen da:

$$X<k \to Y=X$$
$$X\geq k \to Y=k$$

Y aldagaiari buruz zorizko mostra bat jasotzen dugula: $y_1, y_2, ..., y_n$; k parametroaren egiantz handieneko zenbateslea eratu behar da.

k parametroaren espazio parametrikoa hau izango da:

$$k\geq \max\{x_1,x_2,\ldots,x_n\}=u_n$$

Egiantz funtziotik abiatzen gara, u_1, u_2, ..., u_n mostraren datu ordenatuak izanik:

$$
\begin{gathered}
L(y_1, y_2, \ldots, y_n; k) = \\
= P[(y_1, y_2, \ldots, y_n); k] = \\
= P[Y = y_1] \cdot P[Y = y_2] \cdot \ldots \cdot P[Y = y_{n-1}] \cdot P[Y = y_n] = \\
= P[Y = u_1] \cdot P[Y = u_2] \cdot \ldots \cdot P[Y = u_{n-1}] \cdot P[Y = u_n] \cdot n! = \\
P[X = u_1] \cdot P[X = u_2] \cdot \ldots \cdot P[X = u_{n-1}] \cdot P[Y = u_n] \cdot n!
\end{gathered}
$$

X aldagaiari dagokion eredua k-ren menpe ez dagoenez, egiantz funtzioa $P[Y=u_n]$ probabilitatearen araberakoa bakarrik izango da. Beraz, egiantz funtzioa maximotzen duen k balioa, $P[Y=u_n]$ maximotzen duen k balioa izango da. Nola kalkulatzen da $P[Y=u_n]$?

$$
P[Y = u_n] = \begin{cases} P[X = u_n], & k > u_n \\ P[X \geq u_n], & k = u_n \end{cases}
$$

Argi ikusten da $P[Y=u_n]$ probabilitatea, eta ondorioz egiantz funtzioa handiagoa izango dela $k=u_n$ denean, $k>u_n$ denean baino, $P[X \geq u_n] > P[X=u_n]$ betetzen baita orohar. Beraz, k parametroaren egiantz handieneko zenbateslea mostran jasotako daturik handiena izango da: $\hat{k}_{EH} = u_n$.

Esan behar da zenbatesle honek k parametroa gutxiesteko joera duela; hau da, alborapen negatiboa du. Intuitiboki azalduko dut hau: u_n zenbateslearen lagin banakuntzak k izango du goi muga. Beraz, orohar bere itxaropena k baino txikiagoa izango da.

Azkenik, azpimarratu behar da zenbatesle hau banaketari buruz askea dela (‘free distribution estimator’), aurretik ez baita inongo eredu jakinik ezarri. Muga minimo baten zenbatespena ere modu berean planteatu behar da: egiantz handieneko zenbateslea kasu honetan jasotako daturik txikiena da.