ESTATISTIKA METODOAK INGENIARITZAN

8. χ^2 Banaketaren Aplikazioak

eman ta zabal zazu
Universidad del País Vasco
Euskal Herriko Unibertsitatea

8. χ^2 Banaketaren Aplikazioak

8.1 Sarrera

Hipotesi-kontrastea

Populazioko parametro bati buruzko erabaki bat hartzean datza. Hipotesi-kontraste parametrikoak deritze.

Hipotesi estatistikoa

Populazioaren ezaugarri bati buruz egiten den baieztapen bat da.

Hipotesi-kontrastea (Hipotesi kontraste parametrikoak)

Populazioaren ezaugarri bati buruz egindako hipotesia onargarria ala errefusagarria den erabakitzeko erabiltzen den tresna da.

Hipotesi-kontraste ez-parametrikoak

Bi faktoreen arteko independentzia, homogeneotasun-probak edo lagin bateko datuei dagokien banaketa aztergai.

χ^2 banaketa aplikatuko da

8.2 Doikuntzaren egokitasuna

X aztergai

Lagina

$x_1, x_2, \dots, x_k$ n elementu askez osaturik

$f_1, f_2, \dots, f_k$ behatutako maiztasunak

$$\sum_{i=1}^{k} f_i = n$$

$p_1, p_2, \dots, p_k$ probabilitate teorikoak non $p_i = P(X = x_i)$

Itxarotako maiztasunak: $e_i = n.p_i \qquad (i = 1,2, \dots, k)$

$$\sum_{i=1}^{k} p_i = 1 \text{ eta } \sum_{i=1}^{k} e_i = n$$

8.2 Doikuntzaren egokitasuna

Aukeratutako banaketa teorikoak behatutakoarekin bat egiten dueneko hipotesi nulua onargarria den ikusteko, probarako hurrengo estatistikoa erabiliko da:

$$\chi^2_{k-1} = \sum_{i=1}^{k} \frac{\left(f_i - e_i\right)^2}{e_i}$$

Horrela,

$$\sum_{i=1}^{k} \frac{\left(f_i - e_i\right)^2}{e_i} \leq \chi^2_{\alpha;k-1} \Rightarrow$$ H_o onartu α adierazgarritasun mailaz

$$\sum_{i=1}^{k} \frac{\left(f_i - e_i\right)^2}{e_i} > \chi^2_{\alpha;k-1} \Rightarrow$$ H_o errefusatu α adierazgarritasun mailaz

Universidad del País Vasco Euskal Herriko Unibertsitatea

8.2 Doikuntzaren egokitasuna

Sarrera

Doikuntzaren egokitasuna

Bi faktoreen arteko independentzia

Homogeneotasun probak

Yates-en zuzenketa

Kalkuluak errazteko asmoz, probarako testaren berdintza aplikatu:

$$\chi^2_{k-1} = \sum_{i=1}^{k} \frac{\left(f_i - e_i\right)^2}{e_i} = \sum_{i=1}^{k} \frac{f_i^{\,2}}{e_i} - n$$

χ^2 –ren testa aplikatzeko, hurrengo baldintzak hartu behar dira kontuan:

1. Itxarotako maiztasun guztiek $e_i = n.p_i > 5 \quad (i = 1,2,\dots,k)$ bete behar dute
2. Itxarotako maiztasunen batek baldintza hori betetzen ez badu, elkarren segidako modalitateak elkartuko dira, itxarotako maiztasun berriak bost zenbakia gainditu arte
3. Itxarotako maiztasunak lortzeko r parametro estimatu behar badira, orduan χ^2 banaketaren askatasun-graduak k-r-1 dira

Universidad del País Vasco Euskal Herriko Unibertsitatea

8.3 Bi faktoreen arteko independentzia-proba

Sarrera

Doikuntzaren egokitasuna

Bi faktoreen arteko independentzia

Homogeneotasun probak

Yates-en zuzenketa

Izan bitez X eta Y faktoreak

X eta Y independenteak diren ikusteko, χ^2 proba erabiliko duen hurrengo hipotesi-kontrastea egingo da:

H_0: X eta Y independenteak dira
H_a. X eta Y ez dira independenteak

Pausoak:

1. Aztergai den populaziotik n tamainako lagina hartu
2. X eta Y faktoreei dagokien kontingentzia-taula eraiki, gelaxka bakoitzean behatutako maiztasunak adieraziz

8.3 Bi faktoreen arteko independentzia-proba

Kontingentzia taula

X \ Y	y_1	…	y_j	…	y_m	X-ren maiztasunak
x_1	f_{11}	…	f_{1j}	…	f_{1m}	f_{x1}
⋮	⋮	⋮	⋮	⋮	⋮	⋮
x_i	f_{i1}	…	f_{ij}	…	f_{im}	f_{xi}
⋮	⋮	⋮	⋮	⋮	⋮	⋮
x_k	f_{k1}	…	f_{kj}	…	f_{km}	f_{xk}
Y-ren maiztasunak	f_{y1}	…	f_{yj}	…	f_{ym}	n

8.3 Bi faktoreen arteko independentzia-proba

Pausoak:

3. X eta Y independenteak direla suposatuz, e_{ij} itxarotako maiztasuna hurrengo eran kalkulatu

$$e_{ij} = \frac{f_{xi} \cdot f_{yj}}{n}, \qquad i = 1, 2, ..., k \quad j = 1, 2, ..., m$$

Non f_{xi} gaia x_i balioari dagokion total marginala eta f_{yj} gaia y_j balioari dagokion total marginala diren

*Itxarotako maiztasun guztiek $e_{ij} > 5 \quad (i = 1,2, \dots, k, j = 1,2, \dots, m)$ bete behar dute. Kontrako kasuan, ondoz ondoko errenkada edo zutabeak bildu behar dira, itxarotako maiztasun guztiek bost zenbakia gainditu arte.

4. Kontingentzia-taulako datuei dagokien probarako estatistikoaren balioa kalkulatuko da.

$$\chi^2_{(k-1)(m-1)} = \sum_{i=1}^{k} \sum_{j=1}^{m} \frac{\left(f_{ij} - e_{ij}\right)^2}{e_{ij}}$$

8.3 Bi faktoreen arteko independentzia-proba

Horrela,

$$\chi^2_{(k-1)(m-1)} = \sum_{i=1}^{k}\sum_{j=1}^{m}\frac{\left(f_{ij}-e_{ij}\right)^2}{e_{ij}} \le \chi^2_{\alpha;(k-1)(m-1)}$$

$\Downarrow$

H_o onartu α adierazgarritasun mailaz

$$\chi^2_{(k-1)(m-1)} = \sum_{i=1}^{k}\sum_{j=1}^{m}\frac{\left(f_{ij}-e_{ij}\right)^2}{e_{ij}} > \chi^2_{\alpha;(k-1)(m-1)}$$

$\Downarrow$

H_o errefusatu α adierazgarritasun mailaz

8.4 Homogeneotasun probak

Bi faktoreen independentzia aztertzeko orduan χ^2 proba aplikatu da, populazioko lagin bakarra hartuta.

Homogeneotasun probetan populazioko zenbait lagin aske desberdin aukeratuko dira, izaera konkretu batekiko populazioko lagin ezberdinen portaera ikusi nahi baita.

Kalkulu matematikoen aldetik, independentziaren proba eta homogeneotasun proba berdinak dira.

8.5 Yates-en zuzenketa

Sarrera

Doikuntzaren egokitasuna

Bi faktoreen arteko independentzia

Homogeneotasun probak

Yates-en zuzenketa

Bi bider bi motako kontingentzia taula

X \ Y	y_1	y_2	X-ren maiztasunak
x_1	f_{11}	f_{12}	f_{x1}
x_2	f_{21}	f_{22}	f_{x2}
Y-ren maiztasunak	f_{y1}	f_{y2}	n

Kasu honetan probarako estatistikoa:

$$\chi_1^2 = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\left(f_{ij}-e_{ij}\right)^2}{e_{ij}}$$

Errore txikiagoa egiten da hurrengoa erabiliz

Yates-en zuzenketa $$\chi_Y^2 = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\left(\left|f_{ij}-e_{ij}\right|-\frac{1}{2}\right)^2}{e_{ij}}$$

Egokia da balidn eta: $\left|f_{11}\cdot f_{22}-f_{12}\cdot f_{21}\right| > \frac{n}{2}$

eman ta zabal zazu
Universidad del País Vasco Euskal Herriko Unibertsitatea