WIKIPÉDIA

Kappa de Fleiss

Kappa de Fleiss (nommé d'après Joseph L. Fleiss) est une mesure statistique qui évalue la concordance lors de l'assignation qualitative d'objets au sein de catégories pour un certain nombre d'observateurs. Cela contraste avec d'autres kappas tel que le Kappa de Cohen, qui ne fonctionne que pour évaluer la concordance entre deux observateurs. La mesure calcule le degré de concordance de la classification par rapport à ce qui pourrait être attendu si elle était faite au hasard. Il n'y a pas de mesure significative convenue, bien que certaines lignes directrices aient été données.

Le Kappa de Fleiss peut être utilisé classification binaire ou nominale. Il n'y a pas de version disponible pour les classifications de type ordonnée telles que celle de Likert.

Introduction

Le kappa de Fleiss est une généralisation de la statistique du Pi de Scott, Scott, W. (1955)[1], une mesure statistique de concordance[2]. Il est également lié à l'indicateur kappa de Cohen. Mais tandis que le Pi de Scott et le kappa de Cohen fonctionnent pour seulement deux observateurs, le kappa de Fleiss fonctionne pour n'importe quel nombre d'observateurs donnant une classification nominale, à un nombre fixe d'éléments. Il peut être interprété comme exprimant à quel point la valeur de concordance observée parmi les observateurs dépasse ce qui aurait pu être attendu si tous les observateurs avaient fait leur classification totalement au hasard. Il est important de noter que tandis que le kappa de Cohen considère que les deux mêmes observateurs classent un ensemble d'éléments, le kappa de Fleiss considère spécifiquement que bien qu'il y ait un nombre fixe d'observateurs (par exemple 3), différents éléments sont classés par différents individus (Fleiss, 1971, p. 378). Cela signifie que l'élément 1 est classé par les Observateurs A, B et C; mais l'élément 2 pourrait être classé par les Observateurs D, E et F. La concordance peut être considérée comme suit, si un nombre fixe de personnes assigne des classes numériques à un nombre d'éléments, alors le kappa va donner une mesure de la fiabilité de la classification. Kappa, κ , peut être défini comme,

$$\kappa = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e} \qquad (1)$$

Le facteur $1 - \bar{P}_e$ donne le degré de concordance qui est réalisable au-delà du hasard, et, $\bar{P} - \bar{P}_e$ donne le degré de concordance réellement atteint au-dessus du hasard, si les observateurs sont en accord complet, alors $\kappa = 1$. S'il n'y a pas de concordance parmi les observateurs (autre que ce qui aurait pu être atteint par hasard) alors $\kappa \le 0$.

Un exemple d'utilisation du Kappa de Fleiss peut être le suivant: Considérons que 14 psychiatres doivent observer 10 patients. Chaque psychiatre donne un des cinq diagnostics possibles à chacun des patients. Le kappa de Fleiss peut être calculé par la matrice (voir l'exemple ci-dessous) pour montrer le degré de concordance entre les psychiatres au-dessus du niveau de concordance attendu si réalisé au hasard.

Équations

Soit N le nombre total de sujets, soit n le nombre de classification par sujet, et soit k le nombre de catégorie dans laquelle les attributions sont faites. Les sujets sont indexés par i = 1, ... N les catégories sont indexées par j = 1, ... k. Soit n_{ij} qui représente le nombre d'observateurs qui attribuent le i-ème sujet à la j-ème catégorie.

Calculons d'abord p_j, la proportion de toutes les attributions à la j-ème catégorie:

$$p_j = \frac{1}{Nn} \sum_{i=1}^{N} n_{ij}, \qquad 1 = \frac{1}{n} \sum_{j=1}^{k} n_{ij} \qquad (2)$$

Calculons maintenant P_i , qui représente à quel point les observateurs sont d'accord entre eux pour le i-ème sujet (c-à-d, combien d'observateurs--paires d'observateurs sont en accord comparativement au nombre total de l'ensemble des paires observateur-observateur possibles):

$$P_i = \frac{1}{n(n-1)} \sum_{j=1}^{k} n_{ij}(n_{ij} - 1) \qquad (3)$$

$$= \frac{1}{n(n-1)} \sum_{j=1}^{k} (n_{ij}^2 - n_{ij})$$

$$= \frac{1}{n(n-1)} [(\sum_{j=1}^{k} n_{ij}^2) - (n)]$$

Maintenant calculons $\bar{P}$, la moyenne de P_i 's, et $\bar{P}_e$ utilisé dans la formule de κ :

$$\bar{P} = \frac{1}{N} \sum_{i=1}^{N} P_i \tag{4}$$

$$= \frac{1}{Nn(n-1)} (\sum_{i=1}^{N} \sum_{j=1}^{k} n_{ij}^2 - Nn)$$

$$\bar{P}_e = \sum_{j=1}^{k} p_j^2 \tag{5}$$

Exemple d'utilisation

Dans l'exemple suivant, 14 observateurs (n) attribuent à 10 "sujets" (N) un total de 5 catégories (k). Les catégories sont reprises dans les colonnes, tandis que les sujets sont présentés dans les lignes. Chaque case donne le nombre d'observateurs qui sont d'accord sur le fait qu'un sujet appartienne à une certaine catégorie.

Tableau de valeurs pour implémenter l'exemple

	1	2	3	4	5	P_i
1	0	0	0	0	14	1.000
2	0	2	6	4	2	0.253
3	0	0	3	5	6	0.308
4	0	3	9	2	0	0.440
5	2	2	8	1	1	0.330
6	7	7	0	0	0	0.462
7	3	2	6	3	0	0.242
8	2	5	3	2	2	0.176
9	6	5	2	1	0	0.286
10	0	2	2	3	7	0.286
Total	20	28	39	21	32	
p_j	0.143	0.200	0.279	0.150	0.229	

Données

Voir le tableau à droite.

N = 10, n = 14, k = 5

Somme des cases = 140
Somme de P_i = 3.780

Calculs

Par exemple, en prenant la première colonne,

$$p_1 = \frac{0+0+0+0+2+7+3+2+6+0}{140} = 0.143$$

Et en prenant la seconde ligne,

$$P_2 = \frac{1}{14(14-1)} \left(0^2 + 2^2 + 6^2 + 4^2 + 2^2 - 14\right) = 0.253$$

Pour calculer $\bar{P}$, nous devons connaitre P_i,

$$\sum_{i=1}^{N} P_i = 1.000 + 0.253 + \cdots + 0.286 + 0.286 = 3.780$$

À travers l'ensemble de la page,

$$\bar{P} = \frac{1}{(10)} (3.780) = 0.378$$

$$\bar{P}_e = 0.143^2 + 0.200^2 + 0.279^2 + 0.150^2 + 0.229^2 = 0.213$$