

# Finding Derivations: Large

- A contrast-set for which the *maximum difference* between supports is greater than a *minimum support difference threshold*, is called **Large**.

## Example

For the contrast set c1: “*Product = burger  $\wedge$  Store = Fastfood*” and mindev = 5%

support(c1|Inner City Cosmopolitan) = 13%

support(c1|Industrious communities) = 9%

Support(c1|Countryside Living) = 7%

**Deciding if a contrast set is large is straightforward:**

Max difference = 13% - 7% = 6%

With mindev = 5%, c1 is **Large**

To decide if a contrast set is **significant**, we use and statistical test

# Finding Derivations: Significant

- A contrast-set for which at least two groups differ in their support is called **Significant**.
- Perform an statistical test (chi-square) for the contrast set:
- Null hypothesis: “The support for the contrast-set is the same across all groups”
- We build a 2 X k contingency table, k is the number of groups
- Compute the  $\chi^2$  statistics and check value in the chi-square distribution
- Must be less than a threshold  $\alpha$ . (Typically,  $\alpha = 0.05$  )

c1: “*Product = burger  $\wedge$  Store = Fastfood*”

	Inner City	Industrious	Countryside
c1	944	776	172
$\neg$ c1	2745	1851	4597

$$\chi^2 = \sum_{i=1}^2 \sum_{j=1}^k \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \rightarrow \text{Expected values}$$

$$E_{ij} = \frac{\sum_{i=1}^2 O_{ij} \sum_{j=1}^k O_{ij}}{N} \rightarrow \text{Observed values}$$



Consumer  
Data  
Research  
Centre

An ESRC Data  
Investment

KANTAR W<sup>o</sup>RLDPANEL