

## 6. ALTRES CONSIDERACIONS METODOLÒGIQUES

Així mateix, quan el nombre  $n$  dels valors de la variable territorial analitzada sigui prou gran, resultaran poc manejables les taules estadístiques que acullin tots els valors amb les seves corresponents freqüències. En aquests casos, s'agruparan els valors de la variable aleatòria en "classes", que podran ésser de la mateixa o diferent amplitud<sup>9</sup>. Quan esdevingui això, el càlcul de la desviació típica necessària per a la recerca dels CV i dels pertinents coeficients d'uniformitat enregistrarà una mica d'error, degut, precisament, a l'error d'agrupament en classes. Per a ajustar-nos millor a la realitat, s'utilitzarà, llavors, la variància corregida, oferta per l'anomenada "correcció Sheppard", a saber:

$$\sigma_c^2 = \sigma^2 - C^2/12$$

essent  $C$  l'amplitud de l'interval de classe escollit i  $\sigma^2$  la variància de les dades agrupades, i això tindrà lloc en distribucions contínues on les "cues" van gradualment a 0 en ambdues direccions.

En el cas que ens ocupa del Cens Agrari de Catalunya, els sis intervals de classe emprats ho són de diferent amplitud, per raó de la pròpia naturalesa de la dimensió i el nombre d'explotacions.

En línies generals, veurem que un nombre excessiu de "classes" redueix els avantatges de l'agrupació, però un nombre escàs d'intervals pot fins i tot anul·lar la significació de les dades del problema. Respecte a l'amplitud de les "classes" establertes, convé observar que, en general, és convenient que sigui la mateixa per a totes; tanmateix, això dependrà molt de les pròpies dades i de l'objectiu final de la distribució territorial en estudi. En principi, si la distribució és prou uniforme, totes les "classes" seran d'igual amplitud, i si, pel contrari, presenta grans oscil·lacions, pot ésser interessant considerar intervals d'amplitud diferent, com és el cas del que succeeix al nostre estudi (FRANQUET, 1991).

Veurem a la fi, que en base als mateixos o pareguts conceptes, fóra possible la definició d'altres coeficients d'uniformitat territorial. I així, valgui com a exemple el que tindrà en compte el valor del 1r i 3r quartil de la distribució de freqüències de la variable territorial contemplada, a saber:

---

<sup>9</sup> Una norma pràctica genèrica podrà ésser la d'establir una mateixa amplitud equivalent, aproximadament, al 10% de l'observació major, amb la qual cosa el nombre de classes oscil·larà als voltants de la desena. Un altre criteri emprat amb sovintesa és que el nombre de classes s'apropi a l'arrel quadrada del nombre de dades, com ja hem assenyalat. També la Regla de Sturges constitueix una recomanació respecte del nombre desitjable de classes que s'ha de considerar en l'elaboració d'un histograma, tot seguint l'expressió:

$$k = \text{nombre de classes} = 1 + 3 \cdot 3 \cdot \log_{10} n$$

essent  $n$  la grandària del col·lectiu analitzat.

$$CU_5 = 100 \times \sqrt{\frac{Q_1}{Q_3}}$$

que, en el cas d'una distribució moderadament asimètrica (aproximadament normal), oferirà:

$$Q_3 - Q_1 \approx 4\sigma/3 \quad ; \text{ això és:}$$

$$(Q_3 - Q_1)/Q_3 \approx 4\sigma/3Q_3 \approx 1 - Q_1/Q_3 \quad ; \text{ d'on:}$$

$$\sqrt{\frac{Q_1}{Q_3}} = \sqrt{1 - \frac{4\sigma}{3Q_3}}, \text{ amb la qual cosa, també:}$$

$$CU_5 = 100 \times \sqrt{\frac{Q_1}{Q_3}} = \sqrt{10.000} \times \sqrt{1 - \frac{4\sigma}{3Q_3}} = \sqrt{10.000 - \frac{40.000\sigma}{3Q_3}}$$

