

Statistical models

Statistical models are generally used in the frame of a deductive approach. The question concerns variability of a particular phenomenon, and it is attempted to grasp components of this variability. If the statistic individuals are spatial entities, it is the same as wondering about a spatial differentiation. If the statistic individuals are firms, the question concerns differentiations between firms, etc. The approach consists in elaborating a number of hypotheses about the causes of this differentiation and these hypotheses are then tested by the statistic model. The most classical model is multiple regression which applies when all concerned variables are quantitative. The model is formulated:

$$Y = a_1 X_1 + a_2 X_2 + \dots + a_p X_p + e$$

where Y designates the variable " to explain "
X1, X2,.....Xp designate the " explanatory " variables
a1, a2,.....ap are the coefficients
e is the residual of regression.

Application of the model allows, by way of inferential statistics, to estimate the significance and respective roles of X1, X2,...Xp in " explaining " the variability of the phenomenon described by Y. For example, if Y represents the average annual variability ratio of the population of cities in a country, " explicative " variables could be: initial size of a city, qualification of the workforce (part of executives in the working population for example), economic profile (part of active workers occupied in different activity sectors), relative situation (distance to the closest neighbouring city, number of cities in the neighbourhood).

Qualitative explicative variables may be introduced by way of an appropriate coding. To the previous example may thus be added the presence of a HST station or of an airport, etc.

On the opposite, if the variable " to explain " is qualitative, the frame of multiple regression is not appropriate any more. The most frequently used model is then the logit model which expresses the probability to observe such or such modality of the variable to explain in function of explicative, qualitative and possibly quantitative, variables.

Bibliographie