

Modelos estadísticos

Los modelos estadísticos se utilizan en general en el marco de una marcha deductiva. La cuestión alude a la variabilidad de un fenómeno particular, y se busca comprender los componentes de esta *variabilidad*. Si los individuos estadísticos son entidades espaciales, se impone cuestionarse acerca de una diferenciación espacial. Si los individuos estadísticos son empresas, cabe plantear la cuestión sobre las diferenciaciones interempresas, etc. La marcha consiste en elaborar un cierto número de hipótesis sobre las causas de esta diferenciación y estas hipótesis son luego verificadas o no según el modelo estadístico. El modelo más clásico es la regresión múltiple, que se aplica cuando el conjunto de variables en juego son cuantitativas. El modelo se expresa del siguiente modo:

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

donde Y designa la variable "a explicar"

X1, X2,... Xp designan las variables "explicativas"

a1, a2,... ap son los coeficientes

e es el residuo de la regresión.

La aplicación del modelo permite, con la ayuda de la estadística inferencial, evaluar el grado de significación y los roles respectivos de X1, X2,... Xp, para "explicar" la variabilidad del fenómeno descrito por Y. Por ejemplo, si Y representa la tasa de variación media anual de la población de las ciudades de un país, las variables "explicativas" podrían ser: el tamaño inicial de la ciudad, la calificación de la mano de obra (la proporción de técnicos en la población activa, por ejemplo), el perfil económico (la proporción de activos que trabajan en diferentes sectores de actividad), la situación relativa (distancia a la ciudad vecina más próxima, número de ciudades que la rodean).

Las variables explicativas cualitativas pueden ser introducidas por medio de un código apropiado. En el ejemplo precedente se podrían entonces agregar variables de accesibilidad, como la presencia de una estación de trenes de gran velocidad (TGV) o de un aeropuerto, etc.

En contrapartida, si la variable "a explicar" es cualitativa, el marco de la regresión múltiple ya no es apropiado. El modelo utilizado más corrientemente es el modelo logit, que expresa la probabilidad de observar tal modalidad de la variable a explicar en función de las variables explicativas, cualitativas y eventualmente cuantitativas.

Bibliographie