

pecificidad del análisis. Su finalidad es comprobar la existencia de diferencias grupales respecto a una única *variable dependiente* (métrica). Para ello se manipula una *variable independiente*, en función de cuyos valores se forman distintos grupos de tratamiento. Constituidos los grupos, se comprueba la media de cada uno de ellos respecto a la *variable dependiente*. Si se observan diferencias entre las *medias grupales*, se procede después a la comparación de las *varianzas grupales*, y a la medición de su *significatividad*.

Interesa que la *varianza* entre los grupos supere a la *varianza intragrupal*. Ello expresa una mayor heterogeneidad entre los grupos, frente a una escasa variabilidad dentro de ellos. Por lo que podría afirmarse la existencia de diferencias entre los grupos.

La *significatividad* de las diferencias grupales se comprueba mediante los estadísticos "t" (si únicamente se han formado dos grupos de tratamiento), y "F" (si son más de dos los grupos creados).

En ésta, como en cualquier prueba de *significatividad*, se comparan los valores empíricos ("t" y "F") con los *teóricos* (mostrados en las tablas de la "t" de Student y de la "F" de Fisher, correspondientes). El proceso es similar al descrito en el *test de la Chi-Cuadrado*. Se fija el *nivel de significatividad*, en función de la precisión que el investigador desee para su estimación (.05, generalmente); y los *grados de libertad* (ahora determinados por el *tamaño muestral* y el número de *variables independientes* consideradas). Siempre que el *valor empírico* supere al *teórico*, las diferencias grupales observadas en la *muestra* adquirirán *significatividad estadística*. Podrán, por tanto, hacerse extensibles al *universo* del que se extrajo la *muestra* (en los niveles de probabilidad fijados).

#### • Regresión simple

Constituye otra *técnica de dependencia* en la que se analiza la relación entre una única *variable independiente* (métrica o no métrica) y una *variable dependiente* (métrica). Pero, a diferencia de la *técnica analítica anterior*, la *finalidad del análisis es la predicción del valor de la variable dependiente a partir del conocimiento de la independiente*. Se cuantifica la relación existente entre ambas variables; y, se establece el *grado de confianza o significatividad de la estimación efectuada*.

La *correlación* entre las dos variables (*dependiente e independiente*) se mide mediante el *coeficiente R de Pearson*. Éste expresa el grado de *covariación* entre las variables, según se aproxime a "0" (inexistencia de asociación) o a "1" (asociación perfecta). También informa de la dirección de la asociación: creciente (si el signo es positivo) o decreciente (si es negativo).

En el *análisis de regresión*, la idea que subyace es la consecución de una *recta de regresión* que presente el mejor "ajuste" de los casos respecto a las *variables analizadas*. Esta *recta* tiene su expresión matemática en la siguiente *ecuación de regresión*.

$$y = a + bx + e$$

SOC 21-11  
1 COP

donde: "y" denota el valor de la variable *dependiente*.

"a" es el *intercepto* o punto de la *recta* que corta al eje de las Y.

"b" es la *pendiente de la recta* (también referido como el *coeficiente de regresión*). Su *valor* expresa la *cantidad de variación de la variable dependiente por cada unidad de variación de la independiente*. Su *signo* denota si se produce aumento (pendiente creciente; signo positivo) o disminución (pendiente decreciente; signo negativo).

"e" representa el *error de la estimación*: la *inadecuación de la ecuación de regresión* en la predicción del valor de la variable *dependiente*.

Esta *ecuación* permite la *predicción del valor de la variable dependiente a partir de valores conocidos de la independiente*. Los *coeficientes* se obtienen, generalmente, siguiendo el *criterio de mínimos cuadrados* (hacer mínima la distancia que separa los puntos —obtenidos de la confluencia de ambas variables en cada uno de los casos— y la *recta de regresión*).

La *significatividad* de los *coeficientes* se comprueba mediante el estadístico "t", con *n-1 grados de libertad* (siendo "n" el número de observaciones). En cambio, la *significatividad de la correlación* se comprueba mediante el estadístico "F". Como en cualquier prueba de *significatividad*, los valores de "t" y de "F" empíricos han de superar los *teóricos* (determinados en las tablas respectivas) para que el modelo de *regresión* sea *significativo estadísticamente*.

#### 9.4. El análisis multivariable

Los *análisis univariantes y bivariantes* con frecuencia se muestran insuficientes para cubrir los objetivos de la investigación. El proporcionar una visión conjunta e integrada, que describa y/o explique la realidad que se analiza, demanda la realización de *análisis multivariantes* (de más de dos variables al mismo tiempo). De otra forma no podrían medirse las influencias e interrelaciones existentes entre grupos de variables. Como hace tiempo reconociera García Ferrando (1979: 198):

"Las distribuciones bivariantes en sociología aparecen demasiado simplistas para lograr adecuadas explicaciones científicas."

La peculiaridad del *análisis multivariable* reside en operar con un número elevado de variables, y de manera simultánea, basándose en el cálculo matricial. Kendall (1975) lo define como el conjunto de técnicas estadísticas que permite el análisis simultáneo de más de dos variables en una muestra de observaciones. A esta definición, Dillon y Goldstein (1984) añaden la posibilidad de análisis sincrónicos de mediciones en más de una muestra.