

## Analisis Korelasi Kanonik

Analisis korelasi digunakan untuk mengukur kekuatan hubungan antara dua buah variabel, namun kenyataannya kita lebih sering ingin mengetahui bagaimana hubungan antara sekumpulan variabel dengan sekumpulan variabel lainnya, misal kita ingin mengetahui hubungan antara dua set variabel berikut:

- set variabel bakat and set variabel achievement,
- set variabel kepribadian and set variabel yang mengukur kemampuan,
- set variabel penyusun indeks harga set variabel penyusun indeks produksi,
- set variabel ekologis dan set variable lingkungan,

Oleh karena itu, kita tidak bisa menggunakan analisis korelasi untuk mengukurnya, namun dalam kasus ini kita bisa menggunakan **analisis korelasi kanonik**.

Analisis korelasi kanonik adalah suatu alat statistik untuk mengidentifikasi dan mengkuantifikasi hubungan antara dua jenis kelompok (set) variabel. Analisis korelasi kanonik fokus kepada korelasi antara kombinasi linier variabel-variabel di suatu kelompok dengan kombinasi linier variabel-variabel pada satu kelompok lainnya.

Jika kita ingin mengetahui, hubungan atau korelasi antara  $y$  dan  $x$ , dimana  $y$  adalah kombinasi linier dari variabel-variabel ( $y_1, y_2, \dots, y_p$ ) dan  $x$  adalah kombinasi linier dari variabel-variabel ( $x_1, x_2, \dots, x_q$ ), maka pasangan  $y$  dan  $x$  disebut sebagai variabel kanonik. Misal  $U$  adalah kombinasi linier dari  $y$ ,  $U = a'y$ , dan  $V$  adalah kombinasi linier dari  $x$ ,  $V = b'x$ , maka  $U$  dan  $V$  disebut sebagai variat kanonik dan korelasi antar keduanya disebut sebagai korelasi kanonik.

Korelasi kanonik pada prinsipnya adalah perluasan dari korelasi berganda, yang digunakan dalam analisis regresi berganda, dimana korelasi berganda adalah korelasi antara variabel  $y$ , dalam analisis regresi berganda adalah variabel dependen, dengan satu set variabel  $x_1, x_2, \dots, x_q$ , dalam analisis regresi berganda adalah satu set variabel independen.

Karena korelasi kanonik melibatkan dua set variabel, maka korelasi yang terjadi berada dalam  $s$  dimensi, dimana  $s = \min(p, q)$ , sehingga bisa didefinisikan sebagai berikut:

1. Pasangan pertama dari variabel-variabel kanonik, atau pasangan variat kanonik pertama, adalah pasangan kombinasi linier ( $U_1, V_1$ ) yang memiliki unit variance, yang memaksimalkan korelasi;
2. Pasangan variabel-variabel kanonik kedua, atau pasangan variat kanonik kedua, adalah pasangan kombinasi linier ( $U_2, V_2$ ) memiliki unit variance, yang memaksimalkan korelasi antara semua pilihan yang tidak berkorelasi dengan pasangan pertama dari variabel kanonik.

3. Pada tahap ke- $k$ , pasangan variabel-variabel kanonik ke- $k$ , atau pasangan variat kanonik ke  $k$ , adalah pasangan kombinasi linier ( $u_k, v_k$ ) memiliki unit variance, yang memaksimalkan korelasi antara semua pilihan yang tidak berkorelasi dengan  $k-1$  pasangan variabel kanonik.

Dalam analisis korelasi kanonik ada beberapa asumsi yang harus dipenuhi yaitu:

1. Ada hubungan yang bersifat linier (linearitas) antar dua variabel.
2. Perlunya Multivariate Normality untuk menguji signifikansi setiap fungsi kanonik.
3. Tidak ada Multikolinearitas antar anggota kelompok variabel, baik variable dependen maupun variabel independen.