MULTIVARIATE DISTRIBUTIONS

An n-dimensional random vector $x = [x_1, x_2, \ldots, x_n]'$ is an ordered set of n random variables, each of which describes some aspect of a statistical outcome. We write $x \in \mathcal{R}^n$ to signify that x is a point in a real n-dimensional space.

A function $f(x) = f(x_1, x_2, \ldots, x_n)$ that assigns a probability measure to every point in $\mathcal{R}^n$ is called a multivariate p.d.f.

Consider the partitioned vector

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

To conserve space, we prefer to write this as a transposed row vector $x = [x_1', x_2']'$, where $x_1' = [x_1, x_2, \ldots, x_m]'$ and $x_2' = [x_{m+1}, x_{m+2}, \ldots, x_n]'$. (Notice that, in this notation, the subvector x_1 and the leading scalar element x_1 of the vector x are represented by the same symbols. However, these entities can be distinguished by their context.)

The marginal p.d.f, of x_1 is

$$\begin{aligned} f(x_1) &= \int_{x_2} f(x_1, x_2) dx_2 \\ &= \int_{x_n} \cdots \int_{x_{n+1}} f(x_1, x_2, \ldots, x_m, x_{m+1}, \cdots, x_n) dx_{m+1}, \ldots, d_n, \end{aligned}$$

whereas the conditional p.d.f of x_1 given x_2 is

$$f(x_1|x_2) = \frac{f(x)}{f(x_2)} = \frac{f(x_1, x_2)}{f(x_2)}.$$

The expected value of the ith element of x is

$$\begin{aligned} E(x_i) &= \int_x x_i f(x) dx \\ &= \int_{x_n} \cdots \int_{x_i} \cdots \int_{x_1} x_i f(x_1 \ldots, x_i, \ldots, x_n) dx_1, \ldots, d_{x_i}, \ldots, d_{x_1} \\ &= \int_{x_i} x_i f(x_i) dx_i, \end{aligned}$$

where $f(x_i)$ is the marginal distribution of x_i.

The expected value $E(x)$ of the vector $x = [x_1, x_2, \ldots, x_n]'$ is simply the vector containing the expected values of the elements:

$$\begin{aligned} E(x) &= [E(x_1), E(x_2), \ldots, E(x_n)]' \\ &= [\mu_1, \mu_2, \ldots, \mu_n]'. \end{aligned}$$

The variance–covariance matrix or dispersion matrix of x is a matrix $D(x) = \Sigma = [\sigma_{ij}]$ containing the variances and covariances of the elements:

$$D(x) = \begin{bmatrix} V(x_1) & C(x_1, x_2) & \cdots & C(x_1, x_n) \\ C(x_2, x_1) & V(x_2) & \cdots & C(x_2, x_n) \\ \vdots & \vdots & & \vdots \\ C(x_n, x_1) & C(x_n, x_2) & \cdots & V(x_n) \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \ldots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \ldots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \ldots & \sigma_{nn} \end{bmatrix}.$$

The variance–covariance matrix is specified in terms of the vector x by writing

$$\begin{aligned} D(x) &= E\Big\{\big[x - E(x)\big]\big[x - E(x)\big]'\Big\} \\ &= E\left\{ \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \\ \vdots \\ x_n - \mu_n \end{bmatrix} \begin{bmatrix} x_1 - \mu_1 & x_2 - \mu_2 & \cdots & x_n - \mu_n \end{bmatrix} \right\}. \end{aligned}$$

By forming the outer product within the braces, we get the matrix

$$\begin{bmatrix} (x_1 - \mu_1)^2 & (x_1 - \mu_1)(x_2 - \mu_2) & \cdots & (x_1 - \mu_1)(x_n - \mu_n) \\ (x_2 - \mu_2)(x_1 - \mu_1) & (x_2 - \mu_2)^2 & \cdots & (x_2 - \mu_2)(x_n - \mu_n) \\ \vdots & \vdots & & \vdots \\ (x_n - \mu_n)(x_1 - \mu_1) & (x_n - \mu_n)(x_2 - \mu_2) & \cdots & (x_n - \mu_n)^2 \end{bmatrix}.$$

On applying the expectation operator to each of the elements, we get the matrix of variances and covariances.

Quadratic products of the dispersion matrix. The summation vector $\iota = [1, 1, \ldots, 1]'$ is just a column of units. The inner product of the summation vector with a vector $x = [x_1, x_2, \ldots, x_n]'$ of the same order is just the sum of the elements of the latter vector:

$$\iota' x = \begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 + x_2 + \cdots + x_n = \sum_i x_i.$$

The variance of a the sum of the elements of a random vector x is given by

$$V(\iota' x) = \iota' D(x) \iota.$$

To represent this more explicitly, we may write

$$V(\iota' x) = \begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \ldots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \ldots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \ldots & \sigma_{nn} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

There are a variety of ways in which the quadratic products may be represented in scalar notation:

$$\begin{aligned} V(\iota' x) &= \sum_i \sum_j \sigma_{ij} \\ &= \sum_i \sigma_{ii} + \sum_i \sum_j \sigma_{ij} \qquad i \neq j \\ &= \sum_i \sigma_{ii} + 2 \sum_i \sum_j \sigma_{ij} \qquad i < j. \end{aligned}$$

To generalise this result, let $x = [x_1, x_2, \ldots, x_n]'$ be a vector of random variables and $a = [a_1, a_2, \ldots, a_n]'$ be a vector of constants. Then, the variance of the weighted sum $a'x = \sum_i a_i x_i$ is given $V(a'x) = a'D(x)a$. To see this, consider

$$\begin{aligned} a'D(x)a &= a'E\Big\{\big[x - E(x)\big]\big[x - E(x)\big]'\Big\} \\ &= E\Big\{\big[a'x - E(a'x)\big]\big[ax - E(ax)\big]'\Big\} \\ &= E\Big\{\big[a'x - E(a'x)\big]^2\Big\} = V(a'x). \end{aligned}$$

This also demonstrates that $D(x) = \Sigma = [\sigma_{ij}]$ is a positive semi-definite matrix with the property that, for any vector a of the appropriate order, there is $a'\Sigma a \geq 0$, since $D(a'x) = V(a'x) \geq 0$ on account of the non-negativity of all variances. In scalar notation, there is

$$\begin{aligned} V(a'x) &= V\left(\sum_i a_i x_i\right) \\ &= \sum_i a_i^2 V(x_i) + \sum_i \sum_j a_i a_j C(x_i, x_j) \qquad i \neq j \\ &= \sum_i a_i^2 V(x_i) + 2\sum_i \sum_j a_i a_j C(x_i, x_j) \qquad i < j. \end{aligned}$$