The Covariance

Gene Quinn

Definition

Suppose X_1 and X_2 are jointly distributed random variables with

$$\mathsf{E}(X_1) = \mu_1 \quad \text{and} \quad \mathsf{E}(X_2) = \mu_2$$

Definition

Suppose X_1 and X_2 are jointly distributed random variables with

$$\mathsf{E}(X_1) = \mu_1 \quad \text{and} \quad \mathsf{E}(X_2) = \mu_2$$

Define the **covariance** of X_1 and X_2, denoted by

$$\sigma_{12} \quad \text{or} \quad \mathsf{Cov}(X_1, X_2)$$

as:

$$\mathsf{Cov}(X_1, X_2) \; = \; \sigma_{12} \; = \; \mathsf{E}(X_1 X_2) - \mathsf{E}(X_1)\mathsf{E}(X_2)$$

Definition

Following the usual procedure, we calculate $\mathsf{E}(XY)$ as:

$$\mathsf{E}(XY) = \int\int_S xy\, f_X Y(x, y)\, dx\, dy$$

where S is the region of support of f_{XY}

Definition

A standard way of encoding the relationship between two jointly distributed random variables is the **variance-covariance matrix**.

Definition

A standard way of encoding the relationship between two jointly distributed random variables is the **variance-covariance matrix**.

Definition: The variance-covariance matrix for two jointly distributed random variables X_1, X_2 is the matrix V with:

- The i^{th} diagonal element is $\sigma_i^2 \ = \ \text{Var}(X_i)$
- The off diagonal elements are $\sigma_{ij}, \ i \neq j \ = \ \text{Cov}(X_i, X_j)$

Definition

A standard way of encoding the relationship between two jointly distributed random variables is the **variance-covariance matrix**.

Definition: The variance-covariance matrix for two jointly distributed random variables X_1, X_2 is the matrix V with:

- The i^{th} diagonal element is $\sigma_i^2 \;=\;$ Var(X_i)
- The off diagonal elements are $\sigma_{ij},\; i \neq j \;=\;$ Cov(X_i, X_j)

For a bivariate distribution,

$$V \;=\; \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$

Definition

For a trivariate distribution,

$$V = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{bmatrix}$$

Definition

For a trivariate distribution,

$$V = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{bmatrix}$$

With four variates,

$$V = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 & \sigma_{34} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_4^2 \end{bmatrix}$$

Definition

If the variates are independent, all covariances are zero and V is diagonal:

$$V = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix}$$

Linear Combinations

Suppose $\vec{X}$ is a vector of random variables and $\vec{a}$ is a vector of weights.

Linear Combinations

Suppose $\vec{X}$ is a vector of random variables and $\vec{a}$ is a vector of weights.

The variance of the linear combination

$$\vec{a} \cdot \vec{X}$$

is the quadratic form:

$$\vec{a}' V \vec{a}$$

where a' is represents the (column) vector a written as a row vector.