

457.212 Statistics for Civil & Environmental Engineers

In-Class Material: Class 17

Mathematical Expectations of Linear Functions (A&T: 4.3; Supplement #7)

Linear Function: $Y_k = a_{k,0} + \sum_{i=1}^n a_{k,i} X_i$, $k = 1, \dots, m$

Given: (), () and ()

Want to know: (), (), () and ()

Useful when we are interested in the moments of the outputs only and/or the moments of the inputs are only available (not the type of the distribution).

e.g. Seismic loss of buildings in a region: $L = \sum L_i$

Lifecycle cost of an infrastructure: $C = c_i + r \cdot C_d$

1. Linear Function of a Single Random Variable

$$Y = a_0 + a_1 \cdot X$$

(a) Mean of Y : $\mu_Y = a_0 + a_1 \mu_X$

(b) Variance of Y : $\sigma_Y^2 = a_1^2 \sigma_X^2$

(c) Standard deviation of Y : $\sigma_Y = |a_1| \cdot \sigma_X$

Example 1: Consider a building whose initial construction cost (c_i) is estimated as \$100M and the discount rate (r) is 0.8. The uncertain damage cost (C_d) is a random variable whose mean and standard deviation are \$40M and \$10M, respectively. The mean and standard deviation of the lifecycle cost $C = c_i + r \cdot C_d$.

2. Linear Functions of Multiple Random Variables

$$Y_k = a_{k,0} + \sum_{i=1}^n a_{k,i} X_i, \quad k = 1, \dots, m$$

which means,

$$\begin{aligned} Y_1 &= a_{1,0} + a_{1,1}X_1 + a_{1,2}X_2 + \dots + a_{1,n}X_n \\ Y_2 &= a_{2,0} + a_{2,1}X_1 + a_{2,2}X_2 + \dots + a_{2,n}X_n \\ &\vdots \\ Y_k &= a_{k,0} + a_{k,1}X_1 + a_{k,2}X_2 + \dots + a_{k,n}X_n \\ &\vdots \\ Y_m &= a_{m,0} + a_{m,1}X_1 + a_{m,2}X_2 + \dots + a_{m,n}X_n \end{aligned}$$

(a) Algebraic formula ($n \leq 2$)

- Mean: $\mu_{Y_k} = a_{k,0} + \sum_{i=1}^n a_{k,i} \mu_{X_i}$

- Variance:

$$\begin{aligned} \sigma_{Y_k}^2 &= \sum_{i=1}^n a_{k,i}^2 \sigma_{X_i}^2 + \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n a_{k,i} a_{k,j} \sigma_{X_i} \sigma_{X_j} \rho_{X_i X_j} \\ &= \sum_{i=1}^n a_{k,i}^2 \sigma_{X_i}^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n a_{k,i} a_{k,j} \sigma_{X_i} \sigma_{X_j} \rho_{X_i X_j} \end{aligned}$$

- Covariance:

$$\text{Cov}[Y_k, Y_l] = \sum_{i=1}^n a_{k,i} a_{l,i} \sigma_{X_i}^2 + \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n a_{k,i} a_{l,j} \sigma_{X_i} \sigma_{X_j} \rho_{X_i X_j}$$

Example 2: Consider two functions

$$Y_1 = 2 + 3X_1 + 4X_2$$

$$Y_2 = 1 - X_1 + 2X_2$$

The means of X_1 and X_2 are 20 and 30, respectively. The coefficient of variation of each random variable is 10%. The correlation coefficient between X_1 and X_2 is 0.3.

(a) The means of Y_1 and Y_2 :

(b) The standard deviations of Y_1 and Y_2 :

(c) The covariance between Y_1 and Y_2 :

(d) The correlation between Y_1 and Y_2 :

(b) Matrix formula ($n \geq 3$)

The linear functions in (a) can be given in the following matrix form:

$$\mathbf{Y} = \mathbf{a}_0 + \mathbf{A}\mathbf{X}$$

in which

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{bmatrix}, \mathbf{a}_0 = \begin{bmatrix} a_{1,0} \\ a_{2,0} \\ \vdots \\ a_{m,0} \end{bmatrix}, \mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \text{ and } \mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$$

Given: Mean vector and covariance matrix of \mathbf{X}

$$\mathbf{M}_{\mathbf{X}} = \begin{bmatrix} \mu_{X_1} \\ \mu_{X_2} \\ \vdots \\ \mu_{X_n} \end{bmatrix} \text{ and } \Sigma_{\mathbf{XX}} = \begin{bmatrix} \text{Cov}[X_1, X_1] & \text{Cov}[X_1, X_2] & \cdots & \text{Cov}[X_1, X_n] \\ \text{Cov}[X_2, X_1] & \text{Cov}[X_2, X_2] & \cdots & \text{Cov}[X_2, X_n] \\ \vdots & \vdots & & \vdots \\ \text{Cov}[X_n, X_1] & \text{Cov}[X_n, X_2] & \cdots & \text{Cov}[X_n, X_n] \end{bmatrix}$$

Then,

- Mean:

$$\mathbf{M}_Y = \begin{bmatrix} \mu_{Y_1} \\ \mu_{Y_2} \\ \vdots \\ \mu_{Y_m} \end{bmatrix} = \mathbf{a}_0 + \mathbf{A}\mathbf{M}_X$$

- Covariance matrix:

$$\Sigma_{YY} = \mathbf{A}\Sigma_{XX}\mathbf{A}^T$$

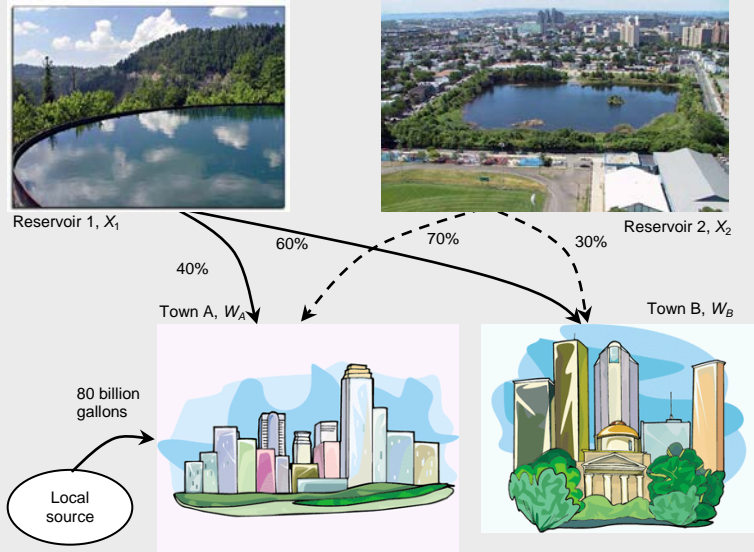
Example 2 (Contd.): Use the matrix formulas for the same problem:

$$Y_1 = 2 + 3X_1 + 4X_2$$

$$Y_2 = 1 - X_1 + 2X_2$$

The means of X_1 and X_2 are 20 and 30, respectively. The coefficient of variation of each random variable is 10%. The correlation coefficient between X_1 and X_2 is 0.3.

Example 3: Consider two towns (A and B) that get water supplies from Reservoir 1, Reservoir 2 and a Local Source. The means of the supplies from Reservoirs 1 and 2, X_1 and X_2 , are 100 and 150 billion gallons, respectively. Their standard deviations are 35 and 30 billion gallons. The correlation coefficient is 0.30.



(a) Describe the water supplies to Town A and B, W_A and W_B in terms of X_1 and X_2 :

(b) The means and standard deviations of W_A and W_B :

(c) The correlation coefficient between W_A and W_B :

(d) Confirm (b) and (c) by use of matrix formulas.

(e) When X_1 and X_2 are normal random variables, $P(W_A < 150)$?