JOINT DISTRIBUTIONS

Outlines

-Discrete/Continuous Random Bivariate Variables

- Joint Probability Distributions

- Marginal Probability Distributions

- Conditional Probability Distributions

- Independence, Covariance and Correlation

-Random vector

Random Vectors

Definition: A random vector is a vector of random variables $\mathbf{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$.

Definition: The mean or expectation of $\mathbf{X}$ is defined as $E[\mathbf{X}] = \begin{pmatrix} E[X_1] \\ \vdots \\ E[X_n] \end{pmatrix}$.

Definition: A random matrix is a matrix of random variables $\mathbf{Z} = (Z_{ij})$. Its expectation is given by $E[\mathbf{Z}] = (E[Z_{ij}])$.

Theorem: A constant vector a (vector of constants) and a constant matrix $\mathbf{A}$ (matrix of constants) satisfy $E[\mathrm{a}] = \mathrm{a}$ and $E[\mathbf{A}] = \mathbf{A}$.

Theorem: $E[\mathbf{X} + \mathbf{Y}] = E[\mathbf{X}] + E[\mathbf{Y}]$.

Theorem: $E[\mathbf{AX}] = \mathbf{A}E[\mathbf{X}]$ for a constant matrix $\mathbf{A}$.

Theorem: $E[\mathbf{AZB} + \mathbf{C}] = \mathbf{A}E[\mathbf{Z}]\mathbf{B} + \mathbf{C}$ if $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are constant matrices.

Definition: If $\mathbf{X}$ is a random vector, the covariance matrix of $\mathbf{X}$ is defined as

$$\text{cov}(\mathbf{X}) \equiv [\text{cov}(X_i, X_j)] \equiv \begin{pmatrix} \text{var}(X_1) & \text{cov}(X_1, X_2) & \cdots & \text{cov}(X_1, X_n) \\ \text{cov}(X_2, X_1) & \text{var}(X_2) & \cdots & \text{cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(X_n, X_1) & \text{cov}(X_n, X_2) & \cdots & \text{var}(X_n) \end{pmatrix}.$$

An alternative form is

$$\text{cov}(\mathbf{X}) = E[(\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])'] = E\left[\begin{pmatrix} X_1 - E[X_1] \\ \vdots \\ X_n - E[X_n] \end{pmatrix} (X_1 - E[X_1], \cdots, X_n - E[X_n])\right].$$

Example: If $X_1, \ldots, X_n$ are independent, then the covariances are 0 and the covariance matrix is equal to $\text{diag}(\sigma_1^2, \ldots, \sigma_n^2)$, or $\sigma^2 \mathbf{I}_n$ if the X_i have common variance σ^2.

Properties of covariance matrices:

Theorem: Symmetry: $\text{cov}(\mathbf{X}) = [\text{cov}(\mathbf{X})]'$.

Theorem: $\text{cov}(\mathbf{X} + \text{a}) = \text{cov}(\mathbf{X})$ if a is a constant vector.

Theorem: $\text{cov}(\mathbf{AX}) = \mathbf{A}\text{cov}(\mathbf{X})\mathbf{A}'$ if $\mathbf{A}$ is a constant matrix.

Theorem: $\text{cov}(\mathbf{X})$ is p.s.d.

Theorem: $\text{cov}(\mathbf{X})$ is p.d. provided no linear combination of the X_i is a constant.

Theorem: $\text{cov}(\mathbf{X}) = E[\mathbf{XX}'] - E[\mathbf{X}](E[\mathbf{X}])'$

Definition: The correlation matrix of $\mathbf{X}$ is defined as

$$\text{corr}(\mathbf{X}) = [\text{corr}(X_i, X_j)] \equiv \begin{pmatrix} 1 & \text{corr}(X_1, X_2) & \cdots & \text{corr}(X_1, X_n) \\ \text{corr}(X_2, X_1) & 1 & \cdots & \text{corr}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{corr}(X_n, X_1) & \text{corr}(X_n, X_2) & \cdots & 1 \end{pmatrix}$$

Definition: If $\mathbf{X}_{m\times 1}$ and $\mathbf{Y}_{n\times 1}$ are random vectors,

$$\mathrm{cov}(\mathbf{X},\mathbf{Y}) = [\mathrm{cov}(X_i, Y_j)] \equiv \begin{pmatrix} \mathrm{cov}(X_1,Y_1) & \mathrm{cov}(X_1,Y_2) & \cdots & \mathrm{cov}(X_1,Y_n) \\ \mathrm{cov}(X_2,Y_1) & \mathrm{cov}(X_2,Y_2) & \cdots & \mathrm{cov}(X_2,Y_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}(X_m,Y_1) & \mathrm{cov}(X_m,Y_2) & \cdots & \mathrm{cov}(X_m,Y_n) \end{pmatrix}.$$

An alternative form is:

$$\mathrm{cov}(\mathbf{X},\mathbf{Y}) = E[(\mathbf{X}-E[\mathbf{X}])(\mathbf{Y}-E[\mathbf{Y}])'] = E\left[\begin{pmatrix} X_1 - E[X_1] \\ \vdots \\ X_m - E[X_m] \end{pmatrix} (Y_1 - E[Y_1], \cdots, Y_n - E[Y_n])\right].$$

Theorem: If $\mathbf{A}$ and $\mathbf{B}$ are constant matrices, then $\mathrm{cov}(\mathbf{AX},\mathbf{BY}) = \mathbf{A}\,\mathrm{cov}(\mathbf{X},\mathbf{Y})\,\mathbf{B}'$.

Theorem: Let $\mathbf{Z} = \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix}$. Then $\mathrm{cov}(\mathbf{Z}) = \begin{pmatrix} \mathrm{cov}(\mathbf{X}) & \mathrm{cov}(\mathbf{X},\mathbf{Y}) \\ \mathrm{cov}(\mathbf{Y},\mathbf{X}) & \mathrm{cov}(\mathbf{Y}) \end{pmatrix}$.

Exercise 1

Let X be a 2×1 discrete random vector and denote its components by X_1 and X_2. Let the support of X be the set of all 2×1 vectors such that their entries belong to the set of the first three natural numbers, i.e.:

$$R_X = \left\{x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top : x_1 \in N_3 \text{ and } x_2 \in N_3\right\}$$

where $N_3 = \{1, 2, 3\}$

Let the joint probability mass function of X be: $p_X(x_1, x_2) = \begin{cases} \frac{1}{36} x_1 x_2 & \text{if } \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top \in R_X \\ 0 & \text{if } \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top \notin R_X \end{cases}$

Find $\mathrm{P}(X_1 = 2 \text{ and } X_2 = 3)$.

Solution

Trivially, we need to evaluate the joint probability mass function at the point $(2, 3)$, i.e.:

$$\begin{aligned} \mathrm{P}(X_1 = 2 \text{ and } X_2 = 3) &= p_X(2, 3) \\ &= \frac{1}{36} \cdot 2 \cdot 3 \\ &= \frac{6}{36} = \frac{1}{6} \end{aligned}$$

Exercise 2

Let X be a 2×1 discrete random vector and denote its components by X_1 and X_2. Let the support of X be the set of all 2×1 vectors such that their entries belong to the set of the first three natural numbers, i.e.:

$$R_X = \left\{ x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top : x_1 \in N_3 \text{ and } x_2 \in N_3 \right\}$$

where $N_3 = \{1, 2, 3\}$

Let the joint probability mass function of X be: $p_X(x_1, x_2) = \begin{cases} \frac{1}{36}(x_1 + x_2) & \text{if } \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top \in R_X \\ 0 & \text{if } \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top \notin R_X \end{cases}$

Find $\mathrm{P}(X_1 + X_2 = 3)$.

Solution

There are only two possible cases that give rise to the occurrence $X_1 + X_2 = 3$. These cases are:

$X = \begin{bmatrix} 1 & 2 \end{bmatrix}^\top$ and $X = \begin{bmatrix} 2 & 1 \end{bmatrix}^\top$

Therefore, since these two cases are disjoint events, we can use the additivity of probability:

$$\begin{aligned} \mathrm{P}(X_1 + X_2 = 3) &= \mathrm{P}\left(\left\{ X = \begin{bmatrix} 1 & 2 \end{bmatrix}^\top \right\} \cup \left\{ X = \begin{bmatrix} 2 & 1 \end{bmatrix}^\top \right\} \right) \\ &= \mathrm{P}\left(X = \begin{bmatrix} 1 & 2 \end{bmatrix}^\top \right) + \mathrm{P}\left(X = \begin{bmatrix} 2 & 1 \end{bmatrix}^\top \right) \\ &= \frac{1}{36}(1 + 2) + \frac{1}{36}(2 + 1) \\ &= \frac{6}{36} = \frac{1}{6} \end{aligned}$$

Exercise 3

Let X be a 2×1 discrete random vector and denote its components by X_1 and X_2.

Let the support of X be: $R_X = \{[1\ \ 1]^\top, [2\ \ 0]^\top, [0\ \ 0]^\top\}$

and its joint probability mass function be:

$$p_X(x) = \begin{cases} \frac{1}{3} & \text{if } x = [1\ \ 1]^\top \\ \frac{1}{3} & \text{if } x = [2\ \ 0]^\top \\ \frac{1}{3} & \text{if } x = [0\ \ 0]^\top \\ 0 & \text{otherwise} \end{cases}$$

Derive the marginal probability mass functions of X_1 and X_2.

Solution

The support of X_1 is: $R_{X_1} = \{0, 1, 2\}$

We need to compute the probability of each element of the support of X_1:

$$p_{X_1}(0) = \sum_{\{(x_1,x_2)\in R_X : x_1 = 0\}} p_X(x_1, x_2)$$

$$= p_X(0,0) = \frac{1}{3}$$

$$p_{X_1}(1) = \sum_{\{(x_1,x_2)\in R_X : x_1 = 1\}} p_X(x_1, x_2)$$

$$= p_X(1,1) = \frac{1}{3}$$

$$p_{X_1}(2) = \sum_{\{(x_1,x_2)\in R_X : x_1 = 2\}} p_X(x_1, x_2)$$

$$= p_X(2,0) = \frac{1}{3}$$

Thus, the probability mass function of X_1 is:

$$p_{X_1}(x) = \sum_{\{(x_1,x_2)\in R_X : x_1 = x\}} p_X(x_1, x_2) = \begin{cases} \frac{1}{3} & \text{if } x = 0 \\ \frac{1}{3} & \text{if } x = 1 \\ \frac{1}{3} & \text{if } x = 2 \\ 0 & \text{otherwise} \end{cases}$$

Continuo

The support of X_2 is: $R_{X_2} = \{0, 1\}$

We need to compute the probability of each element of the support of X_2:

$$p_{X_2}(0) = \sum_{\{(x_1,x_2)\in R_X : x_2 = 0\}} p_X(x_1, x_2)$$

$$= p_X(2,0) + p_X(0,0) = \frac{2}{3}$$

$$p_{X_2}(1) = \sum_{\{(x_1,x_2)\in R_X : x_2 = 1\}} p_X(x_1, x_2)$$

$$= p_X(1,1) = \frac{1}{3}$$

Thus, the probability mass function of X_2 is:

$$p_{X_2}(x) = \sum_{\{(x_1,x_2)\in R_X : x_2 = x\}} p_X(x_1, x_2) = \begin{cases} \frac{2}{3} & \text{if } x = 0 \\ \frac{1}{3} & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases}$$

Exercise 4 Let X be a 2×1 random vector such that its two entries X_1 and X_2 have expected values:

$$E[X_1] = 0$$
$$E[X_2] = 2$$

Let A be the following 2×1 constant vector:

$$A = \begin{bmatrix} 1 \\ 7 \end{bmatrix}$$

Let the random vector Y be defined as follows:

$$Y = A + X$$

Solution

$$\begin{aligned} E[Y] &= E[A + X] \\ &= A + E[X] \\ &= \begin{bmatrix} 1 \\ 7 \end{bmatrix} + \begin{bmatrix} E[X_1] \\ E[X_2] \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 7 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 9 \end{bmatrix} \end{aligned}$$

Exercise 5

Let X be a 1×2 random vector such that

$$E[X_1] = E[X_2] = 3$$

where X_1 and X_2 are the two components of X. Let A be the following 2×2 matrix of constants:

$$A = \begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix}$$

Let the random vector Y be defined as follows:

$$Y = XA$$

Then:

$$\begin{aligned} E[Y] &= E[XA] \\ &= E[X]A \\ &= \begin{bmatrix} E[X_1] & E[X_2] \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 3 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 3 \cdot 2 + 3 \cdot 3 & 3 \cdot 0 + 3 \cdot 1 \end{bmatrix} \\ &= \begin{bmatrix} 15 & 3 \end{bmatrix} \end{aligned}$$

Exercise 6

Let X and Y be two random variables, having expected values:

$$\mathrm{E}[X] = \sqrt{2}$$
$$\mathrm{E}[Y] = 1$$

Compute the expected value of the random variable Z defined as follows:

$$Z = \sqrt{2}X + Y$$

Solution

Using the linearity of the expected value operator, we obtain:

$$\begin{aligned}\mathrm{E}[Z] &= \mathrm{E}\left[\sqrt{2}X + Y\right] \\ &= \sqrt{2}\,\mathrm{E}[X] + \mathrm{E}[Y] \\ &= \sqrt{2}\sqrt{2} + 1 \\ &= 2 + 1 \\ &= 3\end{aligned}$$

Let X be a 2×1 random vector such that its two entries X_1 and X_2 have expected values:

$$\mathrm{E}[X_1] = 2$$
$$\mathrm{E}[X_2] = 3$$

Exercise 7

Let A be the following 2×2 matrix of constants:

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

Compute the expected value of the random vector Y defined as follows:

$$Y = AX$$

Solution

The linearity property of the expected value applies also to the multiplication of a constant matrix and a random vector:

$$\begin{aligned} \mathrm{E}[Y] &= \mathrm{E}[AX] \\ &= A\mathrm{E}[X] \\ &= \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathrm{E}[X_1] \\ \mathrm{E}[X_2] \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 1 \cdot 2 + 2 \cdot 3 \\ 0 \cdot 2 + 1 \cdot 3 \end{bmatrix} \\ &= \begin{bmatrix} 8 \\ 3 \end{bmatrix} \end{aligned}$$

Covariance matrix –

Exercise 8

Let X be a 2×1 random vector and denote its components by X_1 and X_2. The covariance matrix of X is:

$$\mathrm{Var}[X] = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}$$

Compute the variance of the random variable Y defined as:

$$Y = 3X_1 + 4X_2$$

Solution

Using a matrix notation, Y can be written as:

$$Y = \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = bX$$

where we have defined:

$$b = \begin{bmatrix} 3 & 4 \end{bmatrix}$$

Therefore, the variance of Y can be computed using the formula for the covariance matrix of a linear transformation:

$$\begin{aligned} \operatorname{Var}[Y] &= \operatorname{Var}[bX] \\ &= b \operatorname{Var}[X] b^{\top} \\ &= \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 4 \cdot 3 + 1 \cdot 4 \\ 1 \cdot 3 + 2 \cdot 4 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 16 \\ 11 \end{bmatrix} \\ &= 3 \cdot 16 + 4 \cdot 11 \\ &= 92 \end{aligned}$$

Exercise 9

Let X be a 3×1 random vector and denote its components by X_1, X_2 and X_3. The covariance matrix of X is:

$$\mathrm{Var}[X] = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Compute the following covariance:

$$\mathrm{Cov}[X_1 + 2X_3, 3X_2]$$

Solution

Using the bilinearity of the covariance operator, we obtain:

$$\begin{aligned} \mathrm{Cov}[X_1 + 2X_3, 3X_2] &= \mathrm{Cov}[X_1, 3X_2] + 2\mathrm{Cov}[X_3, 3X_2] \\ &= 3\mathrm{Cov}[X_1, X_2] + 6\mathrm{Cov}[X_3, X_2] \\ &= 3 \cdot 1 + 6 \cdot 0 = 3 \end{aligned}$$

The same result can be obtained using the formula for the covariance between two linear transformations. Defining

$$a = \begin{bmatrix} 1 & 0 & 2 \end{bmatrix}$$
$$b = \begin{bmatrix} 0 & 3 & 0 \end{bmatrix}$$

we have:

$$\begin{aligned} &\mathrm{Cov}[X_1 + 2X_3, 3X_2] \\ &= \mathrm{Cov}[aX, bX] \\ &= a\mathrm{Var}[X]b^{\top} \end{aligned}$$

$$= \begin{bmatrix} 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 3 \cdot 0 + 1 \cdot 3 + 0 \cdot 0 \\ 1 \cdot 0 + 2 \cdot 3 + 0 \cdot 0 \\ 0 \cdot 0 + 0 \cdot 3 + 1 \cdot 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$$

$$= 1 \cdot 3 + 0 \cdot 6 + 2 \cdot 0 = 3$$

Exercise 10

Let X be a $K \times 1$ random vector whose covariance matrix is equal to the identity matrix:

$$\mathrm{Var}[X] = I$$

Define a new random vector Y as follows:

$$Y = AX$$

where A is a $K \times K$ matrix of constants such that:

$$AA^{\top} = I$$

Derive the covariance matrix of Y.

Solution

Using the formula for the covariance matrix of a linear transformation:

$$\begin{aligned}\mathrm{Var}[Y] &= \mathrm{Var}[AX] \\ &= A\mathrm{Var}[X]A^{\top} \\ &= AIA^{\top} \\ &= AA^{\top} \\ &= I\end{aligned}$$