AMCS 241 - STAT 250

Lecture # 13

Chapter 7 : Multiple RVs

1- Joint Distributions and Independence

→ Definitions

* The joint PMF of n discrete RVs $X_1, X_2, \ldots, X_n$ is

$$P_{X_1 X_2 \ldots X_n}(x_1, x_2, \ldots, x_n) \triangleq \text{Prob}[X_1 = x_1, X_2 = x_2, \ldots, \text{ and } X_n = x_n]$$

* The joint PDF of n continuous RVs is denoted by $f_{X_1 X_2 \dots X_n}(x_1, x_2, \dots, x_n)$ such as the probability of $A \subset \mathbb{R}^n$ is given by

$$\mathrm{Prob}\left[(X_1, X_2, \dots, X_n) \in A\right] = \int_A f_{X_1 X_2 \dots X_n}(x_1, x_2, \dots, x_n)\, dx_1\, dx_2 \dots dx_n$$

→ <u>Marginals</u>

Marginal of X_1 can be obtained from the joint PDF by

$$f_{X_1}(x_1) = \underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{(n-1)\text{-fold integral}} f_{X_1, X_2 \dots X_n}(x_1, x_2, \dots x_n)\, dx_2\, dx_3 \dots dx_3$$

→ Joint CDF

The joint CDF of n RVs $X_1, X_2, \ldots, X_n$ is given by

$$F_{X_1 X_2 \cdots X_n}(x_1, x_2, \ldots, x_n) = \text{Prob}\left[X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n\right]$$

→ Independence

Random variables $X_1, X_2, \ldots, X_n$ are independent if for all $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, we have

→ $F_{X_1 X_2 \ldots X_n}(x_1, x_2, \ldots, x_n) = F_{X_1}(x_1) . F_{X_2}(x_2) \ldots F_{X_n}(x_n)$

Discrete case

→ $P_{X_1 X_2 \ldots X_n}(x_1, x_2, \ldots, x_n) = P_{X_1}(x_1)\, P_{X_2}(x_2) \ldots P_{X_n}(x_n)$

Continuous case

→ $f_{X_1 X_2 \ldots X_n}(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1) . f_{X_2}(x_2) \ldots f_{X_n}(x_n)$

- **Implication of independence**

  - If $X_1, X_2, \dots, X_n$ are independent then

$$E[X_1 X_2 \dots X_n] = E[X_1] \cdot E[X_2] \dots E[X_n]$$

More generally

$$E[h_1(X_1)\, h_2(X_2) \dots h_n(X_n)] = E[h_1(X_1)]\, E[h_2(X_2)] \dots E[h_n(X_n)]$$

- **i.i.d** (independent identically distributed) RVs are RVs that are independent and have exactly the same marginals i.e

$$F_{X_1}(x) = F_{X_2}(x) = \dots F_{X_n}(x)$$

for all $x \in \mathbb{R}$.

7.2 Random Vectors

7.2.A Notations

- When we have n RVs $X_1, X_2, \dots, X_n$, we can put them in a (column) vector $\underline{X}$:

$$\underline{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ \vdots \\ X_n \end{bmatrix}$$

n-dimensional column vector

- The CDF of a random vector $\underline{X}$

$$F_{\underline{X}}(\underline{x}) = F_{X_1, X_2, \ldots X_n}(x_1, x_2, \ldots, x_n)$$

$$= \text{Prob}[X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n]$$

- The joint PDF can also be written as

$$f_{\underline{X}}(\underline{x}) = f_{X_1 X_2, X_n}(x_1, x_2, \ldots, x_n)$$

7.2.B Expectations

- Expected Value Vector or Mean Vector

$$E[\underline{X}] = \begin{bmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_n] \end{bmatrix}$$

- If we have a random matrix

$$[M] = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & - & X_{2n} \\ & & & \\ X_{n1} & X_{n2} & \cdots & X_{nn} \end{bmatrix}_{n \times n} \qquad [M] = [X_{ij}]_{n \times n}$$

→ The mean Matrix

$$E[[M]] = \begin{bmatrix} E[X_{11}] & E(X_{12}) & \cdots & E[X_{1n}] \\ E[X_{21}] & E(X_{22}) & \ddots & \\ \vdots & & \ddots & \\ E[X_{n1}] & \cdots & \cdots & E(X_{nn}) \end{bmatrix}_{n \times n}$$

→ The linearity of Expectation is also valid for random vectors and random matrices.

→ So for example:

① $\underline{Y} = [A]\ \underline{X} + \underline{b}$

with
- $[A]$ is $m \times n$ fixed matrix
- $\underline{b}$ is a fixed m-dimensional vector

$$E[\underline{Y}] = [A]\ E[\underline{X}] + \underline{b}$$

② If $\underline{X}_1, \underline{X}_2, \ldots \underline{X}_k$ are n-dimensional vectors, then we have:

$$E[\alpha_1 \underline{X}_1 + \alpha_2 \underline{X}_2 + \cdots + \alpha_n \underline{X}_n] =$$

$$\alpha_1 E[\underline{X}_1] + \alpha_2 E[\underline{X}_2] + \cdots + \alpha_n E[\underline{X}_n]$$

7-2-C Correlation and Covariance Matrix.

Def

- For a random vector $\underline{X}$, we define the

1. **Correlation Matrix** of $\underline{X}$ as:

$$[R]_{\underline{X}} = E[\underline{X}\ \underline{X}^T] = \begin{bmatrix} E[X_1^2] & E[X_1 X_2] & \cdots & E[X_1 X_n] \\ \vdots & E[X_2^2] & & \\ \vdots & & \ddots & \\ E[X_n X_1] & \cdots & & E[X_n^2] \end{bmatrix}_{n \times n}$$

2. $\boxed{\text{Covariance Matrix}}$ of $\underline{X}$ as

$$[C]_{\underline{X}} = E\left[\left(\underline{X} - E[\underline{X}]\right)\left(\underline{X} - E[\underline{X}]\right)^T\right]$$

$$= [R]_{\underline{X}} - E[\underline{X}] \cdot E[\underline{X}^T]$$

$$= \begin{bmatrix} Var[X_1] & Cov(X_1 X_2) & \cdots & Cov(X_1 X_n) \\ Cov(X_2 X_1) & Var[X_2] & & \vdots \\ \vdots & & \ddots & \vdots \\ Cov[X_n X_1] & \cdots & \cdots & Var[X_n] \end{bmatrix}_{n \times n}$$

Properties

1. The covariance matrix $[C]_{\underline{x}}$ is a symmetric matrix

$$C_{ij} = Cov[X_i, X_j] = Cov[X_j, X_i] = C_{ji}$$

2. $[C]_{\underline{x}}$ can be diagonalized and all eigenvalues of $[C]_{\underline{x}}$ are real

3. $[C]_{\underline{x}}$ is a positive semi-definite matrix (i.e $\underline{b}^T [C]_x \underline{b} \geq 0$ for all vector $\underline{b}$

- As an extension, we can also introduce for random vectors $\underline{X}$ and $\underline{Y}$

- $[R]_{\underline{X}\,\underline{Y}} = E[\underline{X}\,\underline{Y}^T]$ ← cross-correlation matrix

- $[C]_{\underline{X}\,\underline{Y}} = E[(\underline{X} - E[\underline{X}])(\underline{Y} - E[\underline{Y}])^T]$

↑

cross-covariance matrix

7.3 Functions of Random Vectors

A- Set-Up

- Let $\underline{X}$ be an n-dimensional random vector with joint PDF $f_{\underline{X}}(\underline{x})$.

- Let $G : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a continuous and invertible function with continuous partial derivatives and let us denote

$$H = G^{-1}$$

- Let the random vector $\underline{Y}$ be given by

$$\underline{Y} = G\,\underline{X} \quad (\text{i.e } \underline{X} = G^{-1}(\underline{Y}) = H(\underline{Y}))$$

$$\underline{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} H_1(Y_1, Y_2, \ldots, Y_n) \\ H_2(Y_1, Y_2, \ldots Y_n) \\ \vdots \\ H_n(Y_1, Y_2, \ldots, Y_n) \end{bmatrix}_{n \times 1}$$

B_ Main Result: Method of Transformation

The PDF of the mapped vector $\underline{Y}$,

$f_{\underline{Y}}(\underline{y})$ $\left(f_{Y_1 Y_2 \ldots Y_n}(y_1, y_2, \ldots, y_n)\right)$ is

given by

$$f_{\underline{Y}}(\underline{y}) = f_{\underline{X}}\left(H(\underline{y})\right) \cdot |J|$$

where J is the Jacobian of H defined by

$$J = \det \begin{bmatrix} \frac{\partial H_1}{\partial y_1} & \frac{\partial H_1}{\partial y_2} & \cdots & \frac{\partial H_1}{\partial y_n} \\ \frac{\partial H_2}{\partial y_1} & \frac{\partial H_2}{\partial y_2} & \cdots & \frac{\partial H_2}{\partial y_n} \\ \vdots & & & \\ \frac{\partial H_n}{\partial y_1} & \cdots & & \frac{\partial H_n}{\partial y} \end{bmatrix}_{n \times n}$$

and evaluated at $(y_1, y_2, \dots, y_n)$