SMP
Stochastic Modelling and Processes

Lecture 5

Recap Chapter 5

Joint Distributions

Joint Probability Mass Function

Joint Probability Mass Function

The **joint probability mass function** of the discrete random variables X and Y, denoted as $f_{XY}(x, y)$, satisfies

(1) $f_{XY}(x, y) \geq 0$

(2) $\sum_{X} \sum_{Y} f_{XY}(x, y) = 1$

(3) $f_{XY}(x, y) = P(X = x, Y = y)$ (5.1)

Joint Probability Density Function

Joint Probability Density Function

A **joint probability density function** for the continuous random variables X and Y, denoted as $f_{XY}(x, y)$, satisfies the following properties:

(1) $f_{XY}(x, y) \geq 0$ for all x, y

(2) $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)\, dx\, dy = 1$$

(3) For any region R of two-dimensional space,

$$P((X, Y) \in R) = \iint_R f_{XY}(x, y)\, dx\, dy \qquad (5.2)$$

Marginal Probability Distributions (discrete)

- **The individual probability distribution of a random variable is referred to as its marginal probability distribution.**
- **Consider discrete random variables X and Y. To determine $P(X = x)$, we sum $P(X = x, Y = y)$ over all points in the range of (X, Y) for which $X = x$. Subscripts on the probability mass functions distinguish between the random variables.**

Example 5.3 | Marginal Distribution

The marginal probability distribution for is found by summing the probabilities in each column whereas the marginal probability distribution for is found by summing the probabilities in each row.

For example,

$$f_X(3) = P(X = 3) = P(X = 3, Y = 1) + P(X = 3, Y = 2)$$
$$+ P(X = 3, Y = 3) + P(X = 3, Y = 4)$$
$$= 0.25 + 0.2 + 0.05 + 0.05 = 0.55$$

	x = Number of Bars of Signal Strength			
y = Response Time (nearest second)	1	2	3	Marginal Probability Distribution of Y
4	0.15	0.1	0.05	**0.3**
3	0.02	0.1	0.05	**0.17**
2	0.02	0.03	0.2	**0.25**
1	0.01	0.02	0.25	**0.28**
	0.2	**0.25**	**0.55**	
	Marginal Probability Distribution of X			

FIGURE 5.6

Marginal probability distributions of X and Y from Figure 5.1.

Marginal Probability Density Function

Marginal Probability Density Function

If the joint probability density function of random variables X and Y is $f_{XY}(x, y)$, the **marginal probability density functions** of X and Y are

$$f_X(x) = \int f_{XY}(x, y)\, dy \quad \text{and} \quad f_Y(y) = \int f_{XY}(x, y)\, dx \tag{5.3}$$

where the first integral is over all points in the range of (X, Y) for which $X = x$ and the second integral is over all points in the range of (X, Y) for which $Y = y$.

Mean and Variance from a Joint Distribution

Mean and Variance from a Joint Distribution

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{X,Y}(x, y)dydx$$

and

$$V(X) = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_X)^2 f_{X,Y}(x, y)dydx \qquad (5.4)$$

y = Response Time (nearest second)	x = Number of Bars of Signal Strength 1	2	3	Marginal Probability Distribution of Y
4	0.15	0.1	0.05	**0.3**
3	0.02	0.1	0.05	**0.17**
2	0.02	0.03	0.2	**0.25**
1	0.01	0.02	0.25	**0.28**
	0.2	**0.25**	**0.55**	
	Marginal Probability Distribution of X			

FIGURE 5.6

Marginal probability distributions of X and Y from Figure 5.1.

In Figure 5.6, the marginal probability distributions of X and Y are used to obtain the means as

$$E(X) = 1(0.2) + 2(0.25) + 3(0.55) = 2.35$$

$$E(Y) = 1(0.28) + 2(0.25) + 3(0.17) + 4(0.3) = 2.49$$

Conditional Probability Distributions

Conditional Probability Density Function

Given continuous random variables X and Y with joint probability density function $f_{XY}(x, y)$, the **conditional probability density function** of Y given $X = x$ is

$$f_{Y|x}(y) = \frac{f_{XY}(x, y)}{f_X(x)} \quad \text{for} \quad f_X(x) > 0 \tag{5.5}$$

Because the conditional probability density function $f_{Y|x}(y)$ is a probability density function for all y in R_x, the following properties are satisfied:

(1) $f_{Y|x}(y) \geq 0$

(2) $\int f_{Y|x}(y)\, dy = 1$

(3) $P(Y \in B \mid X = x) = \int_B f_{Y|x}(y)\, dy$ for any set B in the range of Y (5.6)

Example 5.6 | Conditional Probability

Let the random variable X denote the time until a computer server connects to your machine (in milliseconds), and let Y denote the time until the server authorizes you as a valid user (in milliseconds). X and Y measure the wait from a common starting point $(x < y)$.

The joint probability density function for X and Y is

$$f_{XY}(x, y) = 6 \times 10^{-6} \exp(-0.001x - 0.002y) \quad \text{for} \quad x < y$$

From Example 5.2, determine the conditional PDF for Y given $X = x$.

$$f_{Y|x}(y) = \frac{f_{XY}(x, y)}{f_x(x)} = \frac{6 \times 10^{-6} e^{-0.001x-0.002y}}{0.003e^{-0.003x}}$$

$$= 0.002e^{0.002x-0.002y} \quad \text{for} \quad 0 < x \text{ and } x < y$$

Now find the probability that Y exceeds 2000 given that X=1500:

$$P(Y > 2000 | X = 1500)$$

$$= \int_{2000}^{\infty} f_{Y|1500}(y)\,dy$$

$$= \int_{2000}^{\infty} 0.002e^{0.002(1500)-0.002y}$$

$$= 0.002e^{3}\left(\left. \frac{e^{-0.002y}}{-0.002} \right|_{2000}^{\infty} \right)$$

$$= 0.002e^{3}\left(\frac{e^{-4}}{0.002} \right) = e^{-1} = 0.368$$

Conditional Mean and Variance

Conditional Mean and Variance

The **conditional mean** of Y given $X = x$, denoted as $E(Y \mid x)$ or $\mu_{Y|x}$, is

$$E(Y \mid x) = \int_y y \, f_{Y|x}(y) \tag{5.7}$$

and the **conditional variance** of Y given $X = x$, denoted as $V(Y \mid x)$ or $\sigma^2_{Y|x}$, is

$$V(Y \mid x) = \int_y (y - \mu_{Y|x})^2 f_{Y|x}(y) = \int_y y^2 f_{Y|x}(y) - \mu^2_{Y|x}$$

Example 5.7 | Conditional Mean and Variance

For the random variables that denote times in Example 5.2, determine the conditional mean for *Y* given that *x* = 1500. The conditional probability density function for *Y* was determined in Example 5.6. Because $f_{y|1500}(y)$ is nonzero for *y* > 1500,

$$E\left(Y \middle| X = 1500\right) = \int_{1500}^{\infty} y(0.002e^{0.002(1500)-0.002y})\,dy = 0.002e^{3}\int_{1500}^{\infty} ye^{-0.002y}\,dy$$

$$= 0.002e^{3}\left[y\frac{e^{-0.002y}}{-0.002}\bigg|_{1500}^{\infty} - \int_{1500}^{\infty}\left(\frac{e^{-0.002y}}{-0.002}\right)dy \right]$$

$$= 0.002e^{3}\left[\frac{1500}{0.002}e^{-3} - \left(\frac{e^{-0.002y}}{(0.002)(0.002)}\bigg|_{1500}^{\infty} \right) \right]$$

$$= 0.002e^{3}\left[\frac{1500}{0.002}e^{-3} + \frac{e^{-3}}{(0.002)(0.002)} \right]$$

$$= 0.002e^{3}\left[\frac{e^{-3}}{0.002}(2000) \right] = 2000$$

Independence

Independence

For random variables X and Y, if any one of the following properties is true, the others are also true, and X and Y are **independent**.

(1) $f_{XY}(x, y) = f_X(x)\, f_Y(y)$ for all x and y

(2) $f_{Y|x}(y) = f_Y(y)$ for all x and y with $f_X(x) > 0$

(3) $f_{X|y}(x) = f_X(x)$ for all x and y with $f_Y(y) > 0$

(4) $P(X \in A, Y \in B) = P(X \in A)\, P(Y \in B)$ for any sets A and B in the range of X and Y, respectively. (5.8)

Covariance and Correlation

- **A common measure of the relationship between two random variables is the covariance**
- **To define the covariance, we need to describe the expected value of a function of two random variables, *h*(*X*, *Y*)**
- **The definition simply extends the one for a function of a single random variable**

Expected Value of a Function of Two Random Variables

$$E[h(X,Y)] = \begin{cases} \sum\sum h(x,y)\, f_{XY}(x,y) & X, Y \text{ discrete} \\ \iint h(x,y)\, f_{XY}(x,y)\, dx\, dy & X, Y \text{ continuous} \end{cases} \tag{5.14}$$

Covariance

The **covariance** between the random variables X and Y, denoted as $\text{cov}(X, Y)$ or σ_{XY}, is

$$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = E(XY) - \mu_X \mu_Y \tag{5.15}$$

Covariance and Correlation

The covariance between X and Y is defined as

$$\mathrm{Cov}(X,Y) = E\big[(X - EX)(Y - EY)\big] = E[XY] - (EX)(EY).$$

Covariance and Correlation

Lemma. **The covariance has the following properties:**

1) $\mathrm{Cov}(X,X) = E[XX] - EXEX = E[X^2] - (EX)^2 = \mathrm{Var}(X).$

2) $X \& Y$ independent:

$$\mathrm{Cov}(X,Y) = E[XY] - EXEY = E[X]E[Y] - EXEY = 0.$$

3) $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$

4) $\mathrm{Cov}(aX,Y) = a\mathrm{Cov}(X,Y) \quad a \in \mathbb{R}$

Covariance and Correlation

5) $\mathrm{Cov}(X + c, Y) = \mathrm{Cov}(X, Y)$

6) $\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$

7) $\mathrm{Cov}(X+Y, Z+W) = \mathrm{Cov}(X, Z)+\mathrm{Cov}(X, W)+\mathrm{Cov}(Y, Z)+\mathrm{Cov}(Y, W)$

$$\begin{aligned}\mathrm{Cov}(2X + Y, 3Z + W) =& 6\mathrm{Cov}(X, Z) + 2\mathrm{Cov}(X, W)+\\ & 3\mathrm{Cov}(Y, Z) + \mathrm{Cov}(Y, W)\end{aligned}$$

Covariance and Correlation

More generally

$$\mathrm{Cov}\left(\sum_{i=1}^{m} a_i X_i, \sum_{j=1}^{n} b_j Y_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j \mathrm{Cov}(X_i, Y_j).$$

Covariance and Correlation

Variance of a sum:

If $Z = X + Y$, then

$$\begin{aligned} \mathrm{Var}(Z) &= \mathrm{Cov}(Z, Z) \\ &= \mathrm{Cov}(X + Y, X + Y) \\ &= \mathrm{Cov}(X, X) + \mathrm{Cov}(X, Y) + \mathrm{Cov}(Y, X) + \mathrm{Cov}(Y, Y) \\ &= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y). \end{aligned}$$

Covariance and Correlation

More generally, for $a, b \in \mathbb{R}$, we conclude:

$$\mathrm{Var}(aX + bY) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y) + 2ab\mathrm{Cov}(X, Y).$$

Covariance and Correlation

Correlation Coefficient:

$$\rho_{XY} = \rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var(X)\ Var(Y)}}} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}.$$

Covariance and Correlation

Correlation

The **correlation** between random variables X and Y, denoted as ρ_{XY}, is

$$\rho_{XY} = \frac{\operatorname{cov}(X, Y)}{\sqrt{V(X)V(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \tag{5.16}$$

For any two random variables X and Y,

$$-1 \leq \rho_{XY} \leq +1 \tag{5.17}$$

If X and Y are independent random variables,

$$\sigma_{XY} = \rho_{XY} = 0 \tag{5.18}$$

Covariance and Correlation

Properties of the correlation coefficient:

1) $-1 \leq \rho(X, Y) \leq 1;$

2) $\rho(aX + b, cY + d) = \rho(X, Y) \text{ for } a, c > 0;$

3) $\rho(X, Y) = 1 \quad \text{if } Y = aX + b \;\; a > 0;$
$\rho(X, Y) = -1 \;\; \text{if } Y = aX + b \;\; a < 0.$

Covariance and Correlation

Definition. Consider two random variables X and Y :

1) If $\rho(X,Y) = 0$, we say that X and Y are uncorrelated.
2) If $\rho(X,Y) > 0$, we say that X and Y are positively correlated.
3) If $\rho(X,Y) < 0$, we say that X and Y are negatively correlated.

Very important!

i.i.d.: Independent and Identically distributed

- We define that for series of random variables that is taken from the same distribution (identically distributed), and are sampled independent of each other, that they are i.i.d.

$$\boxed{\text{i.i.d.} = \text{Independent and Identically distributed}}$$

- i.i.d. is a very important characteristic in stochastic variable processing and statistics