

STA2001 Probability and Statistics (I)

Lecture 12

Tianshi Chen

The Chinese University of Hong Kong, Shenzhen

Review of the last lecture

Key concepts and/or techniques:

- ▶ Mathematical Expectation:

$$E(g(X, Y)) = \sum_{(x,y) \in \bar{S}} g(x, y) f(x, y)$$

- ▶ Covariance and correlation coefficient:

To study the relation between 2 RVs

$$\text{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$$

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}}, \quad \text{Var}(X) > 0, \text{Var}(Y) > 0.$$

- ▶ Interpretation and properties of covariance and correlation coefficient

Review of the last lecture

- ▶ $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$

When $\text{Cov}(X, Y) = 0$, X and Y are uncorrelated.

When $\text{Cov}(X, Y) > 0$, X and Y are positively correlated.

When $\text{Cov}(X, Y) < 0$, X and Y are negatively correlated.

- ▶ Interpretation: Roughly speaking, a positive or negative covariance indicate that the values of $X - E(X)$ and $Y - E(Y)$ obtained in a single experiment “tend” to have the same or the opposite sign respectively.
- ▶ Independence of X and $Y \Rightarrow$ uncorrelation of X and Y , but the converse is in general not true.

Review of the last lecture

Correlation coefficient

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}}$$

- ▶ It is a normalized version of $\text{Cov}(X, Y)$ and in fact $-1 \leq \rho(X, Y) \leq 1$ and the size of $|\rho|$ provides a normalized measure of the extent to which this is true.
- ▶ $\rho = 1$ or $(\rho = -1)$ if and only if there exists a positive. (or negative, respectively) constant c such that

$$Y - E(Y) = c(X - E(X))$$

Review of the last lecture

Key concepts and/or techniques:

► Conditional distribution

Motivation: it is a probability distribution that describes the distribution of probability of events of a RV given the occurrence of a particular event.

For example, the conditional pmf of X given $Y = y$ is

$$g(x|y) = \frac{f(x, y)}{f_Y(y)}, \quad x \in \overline{S_X}(y)$$

provided that $f_Y(y) > 0$.

► Conditional mathematical expectations

The conditional expectation of $g(Y)$ given $X = x$

$$E(g(Y)|X = x) = \sum_{y \in \overline{S_Y}(x)} g(y)h(y|x)$$

Review of the last lecture

[Conditional pmf]

Conditional pmf of X given $Y = y$ is defined by

$$g(x|y) = \frac{f(x, y)}{f_Y(y)}, \quad x \in \overline{S_X}(y)$$

provided that $f_Y(y) > 0$.

Similarly, conditional pmf of Y given that $X = x$ is defined by

$$h(y|x) = \frac{f(x, y)}{f_X(x)}, \quad y \in \overline{S_Y}(x)$$

provided that $f_X(x) > 0$.

Review of the last lecture

Conditional pmf

Conditional pmf is a well-defined pmf

1. $h(y|x) > 0$
2. $\sum_{y \in \overline{S}_Y(x)} h(y|x) = 1$
3. for $A \subseteq \overline{S}_Y(x)$

$$P(Y \in A | X = x) = \sum_{y \in A} h(y|x)$$

Review of the last lecture

[Conditional Mathematical Expectation]

- ▶ Let $g(Y)$ be a function of Y .

Then the conditional expectation of $g(Y)$ given $X = x$

$$E(g(Y)|X = x) = \sum_{y \in \overline{S}_Y(x)} g(y)h(y|x)$$

- ▶ When $g(Y) = Y$, conditional mean

$$E(Y|X = x) = \sum_{y \in \overline{S}_Y(x)} yh(y|x)$$

- ▶ When $g(Y) = [Y - E(Y|X = x)]^2$, conditional variance

$$\text{Var}(Y|X = x) \triangleq E\{[Y - E(Y|X = x)]^2|X = x\}$$

Section 4.4 Bivariate Distribution of Continuous Type

Bivariate Continuous RV

Definition

Let X and Y be two continuous random variables and (X, Y) be a pair of RVs with their range denoted by $\overline{S} \subseteq R^2$. Then (X, Y) or X and Y is said to be a bivariate continuous RV.

Moreover, let $\overline{S}_X \subseteq R$ and $\overline{S}_Y \subseteq R$ denote the range of X and Y , respectively.

$$\overline{S} = \{\text{all possible values of } (X, Y)\}$$

$$\overline{S}_X = \{\text{all possible values of } X\} = \{x | (x, y) \in \overline{S}\}$$

$$\overline{S}_Y = \{\text{all possible values of } Y\} = \{y | (x, y) \in \overline{S}\}$$

Then, it holds that

$$\overline{S} \subseteq \overline{S}_X \times \overline{S}_Y = \{(x, y) | x \in \overline{S}_X, y \in \overline{S}_Y\}$$

Roadmap for bivariate continuous random distributions

To study the bivariate continuous random variable

discrete RV \longrightarrow continuous RV

pmf \longrightarrow pdf

joint pmf \longrightarrow joint pdf

marginal pmf \longrightarrow marginal pdf

conditional pmf \longrightarrow conditional pdf

Mathematical expectations

mean

variance

covariance

correlation coefficient

Joint pdf

Definition

The joint pdf of two continuous RVs X and Y is a function $f(x, y) : \bar{S} \rightarrow (0, \infty)$ with the following properties:

1. $f(x, y) > 0, (x, y) \in \bar{S}$

2. $\iint_{\bar{S}} f(x, y) dx dy = 1$

3.

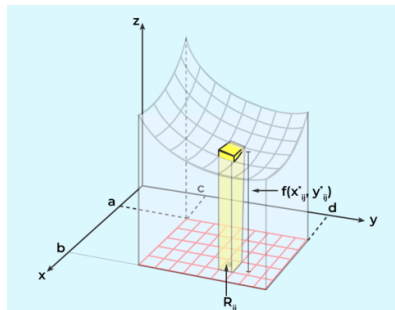
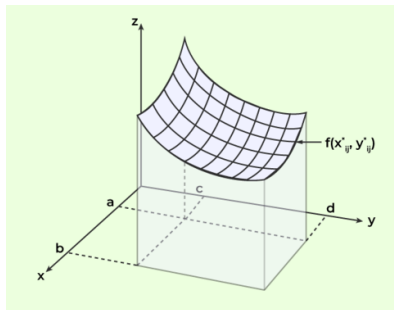
$$\begin{aligned} P((X, Y) \in A) &\triangleq P(\{(X, Y) \in A\}) \\ &= \iint_A f(x, y) dx dy, A \subseteq \bar{S} \end{aligned}$$

Remarks

Recall the geometric interpretation of double integral:

$$P((X, Y) \in A) = \iint_A f(x, y) dx dy$$

calculates the volume of the solid under the surface $z = f(x, y)$ over the region A in the the xy -plane.



Remarks

- ▶ Very often, we extend the definition domain of $f(x, y)$ from \overline{S} to $R \times R$ by letting $f(x, y) = 0$, for $(x, y) \notin \overline{S}$ and $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$

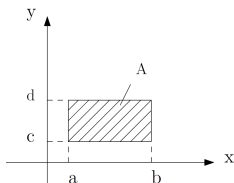
- ▶ Bottom line:

If the set A is rectangular with its line segments parallel to the coordinate axes, i.e.,

$$A = \{(x, y) | a \leq x \leq b, c \leq y \leq d\},$$

then the double integral becomes

$$\begin{aligned} P((X, Y) \in A) &= \int_a^b \int_c^d f(x, y) dy dx \\ &= \int_c^d \int_a^b f(x, y) dx dy \end{aligned}$$



Remarks

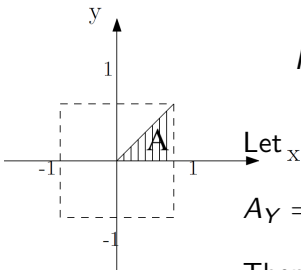
► General case:

Let

$$A_X = \{x | (x, y) \in A\}, A_Y(x) = \{y | (x, y) \in A\} \text{ for } x \in A_X$$

Then

$$P((X, Y) \in A) = \int_{A_X} \int_{A_Y(x)} f(x, y) dy dx$$



$$A_Y = \{y | (x, y) \in A\}, A_X(y) = \{x | (x, y) \in A\} \text{ for } y \in A_Y$$

Then

$$P((X, Y) \in A) = \int_{A_Y} \int_{A_X(y)} f(x, y) dx dy$$

Marginal pdf

Definition

The marginal pdf of X , $f_X(x) : \overline{S_X} \rightarrow (0, \infty)$

$$f_X(x) = \int_{\overline{S_Y}(x)} f(x, y) dy$$

$$\overline{S_Y}(x) = \{y | (x, y) \in \overline{S}\} \text{ for } x \in \overline{S_X}$$

The marginal pdf of Y , $f_Y(y) : \overline{S_Y} \rightarrow (0, \infty)$

$$f_Y(y) = \int_{\overline{S_X}(y)} f(x, y) dx$$

$$\overline{S_X}(y) = \{x | (x, y) \in \overline{S}\} \text{ for } y \in \overline{S_Y}$$

Example 1, page 156

Question

Let X and Y have the joint pdf

$$f(x, y) = \frac{3}{2}x^2(1 - |y|), \quad -1 < x < 1, \quad -1 < y < 1$$

$$\Rightarrow \begin{cases} \overline{S_X} = \{x \mid -1 < x < 1\} \\ \overline{S_Y} = \{y \mid -1 < y < 1\} \\ \overline{S} = \{(x, y) \mid -1 < x < 1, -1 < y < 1\} \end{cases}$$

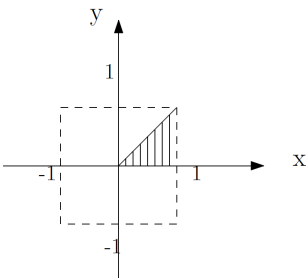
Q1: Let $A = \{(x, y) \mid 0 < x < 1, 0 < y < x\}$. What is the probability of A ?

Q2: What is the marginal pdf of X and Y ?

Q3: What is the expectation of X ?

Example 1, page 156

Q1:



$$\begin{aligned}P(A) &= \int_0^1 \int_0^x \frac{3}{2}x^2(1 - |y|)dydx \\&= \int_0^1 \frac{3}{2}x^2\left(x - \frac{1}{2}x^2\right)dx = \frac{3}{2} \cdot \frac{1}{4}x^4 \Big|_0^1 - \frac{3}{4} \cdot \frac{1}{5}x^5 \Big|_0^1 \\&= \int_0^1 \int_y^1 \frac{3}{2}x^2(1 - |y|)dxdy \\&= \int_0^1 \frac{3}{2} \cdot \frac{1}{3}x^3 \Big|_y^1 (1 - |y|)dy \\&= \frac{9}{40}\end{aligned}$$

Example 1, page 156

Q2:

$$\begin{aligned}\text{For } x \in \overline{S_X}, f_X(x) &= \int_{\overline{S_Y}(x)} f(x, y) dy \\ &= \int_{-1}^1 \frac{3}{2} x^2 (1 - |y|) dy = \frac{3}{2} x^2 (2 + (-1)) = \frac{3}{2} x^2\end{aligned}$$

$$\begin{aligned}\text{For } y \in \overline{S_Y}, f_Y(y) &= \int_{\overline{S_X}(y)} f(x, y) dx \\ &= \int_{-1}^1 \frac{3}{2} x^2 (1 - |y|) dx = \frac{3}{2} (1 - |y|) \frac{1}{3} x^3 \Big|_{-1}^1 \\ &= 1 - |y|\end{aligned}$$

Example 1, page 156

Q3:

$$E(X) = \int_{\overline{S_X}} x f_X(x) dx = \int_{-1}^1 x \frac{3}{2} x^2 dx = 0$$

$$\begin{aligned} E(X) &= \int \int_{\overline{S}} x f(x, y) dx dy = \int_{-1}^1 \int_{-1}^1 \frac{3}{2} x^3 (1 - |y|) dx dy \\ &= \int_{-1}^1 \frac{3}{2} x^3 dx \int_{-1}^1 (1 - |y|) dy = 0 \end{aligned}$$

Mathematical Expectation

Definition

Let $g(X, Y)$ be a function of X and Y , whose joint pdf $f(x, y) : \bar{S} \rightarrow (0, \infty)$. Then

$$E[g(X, Y)] = \iint_{\bar{S}} g(x, y) f(x, y) dx dy$$

► $g(X, Y) = X \rightarrow$ mean of X

$$\begin{aligned} E[X] &= \iint_{\bar{S}} x f(x, y) dx dy \\ &= \int_{\bar{S}_X} x \int_{\bar{S}_Y(x)} f(x, y) dy dx \\ &= \int_{\bar{S}_X} x f_X(x) dx \end{aligned}$$

Mathematical Expectation

Definition

- $g(X, Y) = (X - E[X])^2 \rightarrow$ variance of X

$$\begin{aligned} \text{Var}[X] &= \iint_{\bar{S}} (x - E[X])^2 f(x, y) dx dy \\ &= \int_{\bar{S}_X} (x - E[X])^2 \int_{\bar{S}_Y(x)} f(x, y) dy dx \\ &= \int_{\bar{S}_X} (x - E[X])^2 f_X(x) dx \end{aligned}$$

Example 2, page 155

Question

Let X and Y have the joint pdf

$$f(x, y) = \left(\frac{4}{3}\right)(1 - xy) \quad \text{with} \quad 0 \leq x \leq 1, \quad 0 \leq y \leq 1$$

Q1 Find the marginal pdfs of X and Y ?

Q2 Find the expectation of X ?

Q3 Find the variance of X

Example 2, page 155

Question

Let X and Y have the joint pdf

$$f(x, y) = \left(\frac{4}{3}\right)(1 - xy) \quad \text{with} \quad 0 \leq x \leq 1, \quad 0 \leq y \leq 1$$

Q1 Find the marginal pdfs of X and Y ?

Q2 Find the expectation of X ?

Q3 Find the variance of X

Q1 :

$$\begin{aligned} f_X(x) &= \int_{\overline{S_Y}(x)} f(x, y) dy = \int_0^1 \frac{4}{3}(1 - xy) dy \\ &= \frac{4}{3} - \frac{4}{3}x \frac{1}{2}y^2 \Big|_0^1 = \frac{4}{3}\left(1 - \frac{1}{2}x\right) \end{aligned}$$

Example 2, page 155

$$\begin{aligned}f_Y(y) &= \int_{\overline{S_X}(y)} f(x, y) dx \\&= \int_0^1 \frac{4}{3}(1 - xy) dx = \frac{4}{3}\left(1 - \frac{1}{2}y\right)\end{aligned}$$

Q2:

$$\begin{aligned}E[X] &= \int_{\overline{S_X}} xf_X(x) dx = \int_0^1 x \frac{4}{3}\left(1 - \frac{1}{2}x\right) dx \\&= \frac{4}{3} \cdot \frac{1}{2}x^2 \Big|_0^1 - \frac{4}{6} \cdot \frac{1}{3}x^3 \Big|_0^1 = \frac{4}{9}\end{aligned}$$

Q3:

$$\text{Var}[X] = \int_{\overline{S_X}} (x - E[X])^2 f_X(x) dx = \int_0^1 \left(x - \frac{4}{9}\right)^2 \frac{4}{3}\left(1 - \frac{1}{2}x\right) dx$$

Independent Continuous RVs

Definition

Two continuous RVs X and Y are independent if

$$f(x, y) = f_X(x)f_Y(y), \quad x \in \overline{S_X}, \quad y \in \overline{S_Y}$$

If X and Y are not independent, then we say X and Y are dependent.

When X and Y are independent,

$$\overline{S} = \overline{S_X} \times \overline{S_Y}. \quad \overline{S} \text{ is said to be rectangular}$$

which is a necessary condition for independence of X and Y .

Example 2 — revisited

Note that

$$f(x, y) = \left(\frac{4}{3}\right)(1 - xy) \quad \text{with} \quad 0 \leq x \leq 1, \quad 0 \leq y \leq 1$$

$$f_X(x) = \frac{4}{3}\left(1 - \frac{1}{2}x\right)$$

$$f_Y(y) = \frac{4}{3}\left(1 - \frac{1}{2}y\right)$$

Since $f(x, y) \neq f_X(x)f_Y(y)$, X and Y are NOT independent.

Covariance and Correlation Coefficient

Definition

$$\begin{aligned}\text{Cov}(X, Y) &= E[(X - E[X])(Y - E[Y])] \\ &= E(XY) - E(X)E(Y)\end{aligned}$$

$$E(XY) = \iint_{\bar{S}} xyf(x, y)dx dy.$$

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}}, \quad \text{Var}(X) > 0, \text{Var}(Y) > 0.$$

Conditional pdf

Definition

Let X and Y have a joint pdf $f(x, y) : \bar{S} \rightarrow (0, \infty)$ and marginal pdf $f_X(x) : \bar{S}_X \rightarrow (0, \infty)$ and $f_Y(y) : \bar{S}_Y \rightarrow (0, \infty)$.

The conditional pdf of Y , given that $X = x$ are

$$h(y|x) = \frac{f(x, y)}{f_X(x)} \quad \text{for } f_X(x) > 0, y \in \bar{S}_Y(x)$$

$$\text{For } A \subseteq \bar{S}_Y(x), \quad P(Y \in A | X = x) = \int_{y \in A} h(y|x) dy.$$

Conditional pdf

Definition

Let X and Y have a joint pdf $f(x, y) : \bar{S} \rightarrow (0, \infty)$ and marginal pdf $f_X(x) : \bar{S}_X \rightarrow (0, \infty)$ and $f_Y(y) : \bar{S}_Y \rightarrow (0, \infty)$.

The conditional pdf of Y , given that $X = x$ are

$$h(y|x) = \frac{f(x, y)}{f_X(x)} \quad \text{for } f_X(x) > 0, y \in \bar{S}_Y(x)$$

$$\text{For } A \subseteq \bar{S}_Y(x), \quad P(Y \in A | X = x) = \int_{y \in A} h(y|x) dy.$$

The conditional pdf of X , given that $Y = y$ are

$$g(x|y) = \frac{f(x, y)}{f_Y(y)} \quad \text{for } f_Y(y) > 0, x \in \bar{S}_X(y)$$

$$\text{For } A \subseteq \bar{S}_X(y), \quad P(X \in A | Y = y) = \int_{x \in A} g(x|y) dx.$$

Conditional mathematical expectation

Definition

The conditional mathematical expectation of a function of Y , $g(Y)$, given that $X = x$ is

$$E(g(Y)|X = x) = \int_{\overline{S_Y(x)}} g(y)h(y|x)dy$$

The conditional mean and variance of Y , given that $X = x$ are

$$E(Y|X = x) = \int_{\overline{S_Y(x)}} yh(y|x)dy$$

$$\begin{aligned} \text{Var}(Y|X = x) &= E\{[Y - E(Y|X = x)]^2|X = x\} \\ &= \int_{\overline{S_Y(x)}} [y - E(Y|X = x)]^2 h(y|x)dy \\ &= E[Y^2|X = x] - [E(Y|X = x)]^2 \end{aligned}$$

Example 3, page 157

Question

Let X and Y be two continuous RVs with

$$f(x, y) = 2, \quad 0 \leq x \leq y \leq 1$$

$$\overline{S} = \{(x, y) | 0 \leq x \leq y \leq 1\}, \overline{S_X} = \overline{S_Y} = [0, 1]$$

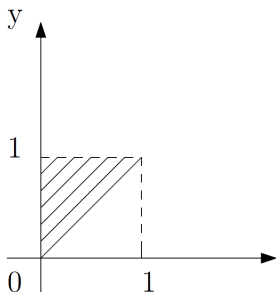
Q1: $f_X(x), f_Y(y), E(X), E(Y)$?

Q2: $h(y|x), E(Y|X = x), \text{Var}(Y|X = x)$?

Q3: $P(\frac{3}{4} \leq Y \leq \frac{7}{8} | X = \frac{1}{4})$?

Example 3, page 157

Q1:



$$f_X(x) = \int_{\overline{S_Y}(x)} f(x, y) dy = \int_x^1 2 dy = 2(1 - x), 0 \leq x \leq 1$$

$$f_Y(y) = \int_{\overline{S_X}(y)} f(x, y) dx = \int_0^y 2 dy = 2y, \quad 0 \leq y \leq 1$$

$$E(X) = \int_{\overline{S_X}} x f_X(x) dx = \int_0^1 x 2(1 - x) = \frac{1}{3}$$

$$E(Y) = \int_{\overline{S_Y}} y f_Y(y) dy = \int_0^1 2y^2 dy = \frac{2}{3}$$

Example 3, page 157

Q2:

$$h(y|x) = \frac{f(x, y)}{f_X(x)} = \frac{1}{1-x}, 0 \leq x \leq y \leq 1.$$

$$\begin{aligned} E(Y|X=x) &= \int_{\overline{S_Y(x)}} y h(y|x) dy = \int_x^1 y \frac{1}{1-x} dy \\ &= \frac{1}{2(1-x)} (1-x^2) = \frac{1}{2}(1+x) \end{aligned}$$

$$\begin{aligned} \text{Var}(Y|X=x) &= \int_{\overline{S_Y(x)}} \left[y - \frac{1}{2}(1+x) \right]^2 h(y|x) dy \\ &= \int_x^1 \frac{1}{1-x} \left[y - \frac{1}{2}(1+x) \right]^2 dy \\ &= \frac{1}{3} \frac{1}{1-x} \left[y - \frac{1}{2}(1+x) \right]^3 \Big|_x^1 = \frac{1}{12} (1-x)^2 \end{aligned}$$

Example 3, page 157

Q3:

$$P\left(\frac{3}{4} \leq Y \leq \frac{7}{8} \mid X = \frac{1}{4}\right) = \int_{\frac{3}{4}}^{\frac{7}{8}} h(y \mid \frac{1}{4}) dy = \frac{4}{3} \left(\frac{7}{8} - \frac{3}{4} \right) = \frac{1}{6}$$