ISyE 3770, Spring 2024
Statistics and Applications

Bivariate Probability Distribution

Instructor: Jie Wang
H. Milton Stewart School of
Industrial and Systems Engineering
Georgia Tech

jwang3163@gatech.edu
Office: ISyE Main 445

Chapter 4 Bivariate Distributions

Section 4.1 Bivariate Distributions with the **discrete** type

➢ Motivation

very often, the outcome of a random experiment is a **tuple** of several things of interests:

- Observe female college students to obtain information such as ***height x***, and ***weight y***.
- Observe high school students to obtain information such as ***rank x***, and ***score of college entrance examination y***.

➢ In order to define joint probability mass function (joint *pmf*):

* **Complete way :**

① identify the Sample Space S;

② Define a *RV* $Z = \begin{bmatrix} X \\ Y \end{bmatrix} : S \to Z(S)$;

③ Define a *pmf* for Z, $f(z): Z(S) \to [0,1]$.

* **Simplified way :**

① Ignore the Sample Space S;

② Specify $Z(S)$ directly and denote it by D;

③ Define the *pmf* for Z, $f(z): D \to [0,1]$;

equivalently, for $\begin{bmatrix} X \\ Y \end{bmatrix}$, $f(x,y): D \to [0,1]$.

Definition [joint probability mass function (joint *pmf*)]

Let X and Y be 2 *RV*s. The probability that $X = x$ and $Y = y$ is denoted by $f(x, y) = P(X = x, Y = y)$.

The function $f(x, y)$: D → [0,1] is called the **joint probability mass function** (**joint *pmf***) of (X,Y) *if*:

① $0 \le f(x,y) \le 1$;

② $\sum_{(x,y)\in D} f(x,y) = 1$;

③ $P[(X,Y)\in A] \triangleq P(\{(x,y)\in A\}) = \sum_{(x,y)\in A} f(x,y), \quad A \subseteq D.$

Example 1

Roll a pair of fair dice. The sample space contains 36 outcomes. And let X denote the smaller outcome and Y the larger outcome on the die.

For instance, if the outcome is (3,2), then X =2, Y =3.

Obviously, $P(\{X=2, Y=3\}) = 1/36 + 1/36 = 2/36.$

$P(\{X=2, Y=2\}) = 1/36.$

Furthermore, the *joint pmf* of X and Y is: $f(x,y) = \begin{cases} 1/36, & 1 \le x = y \le 6 \\ 2/36, & 1 \le x < y \le 6 \end{cases}$

Definition [Marginal *pmf*]

Let X and Y have the joint probability mass function $f(x, y)$: D → [0,1]. Sometimes we are interested in the pmf of X *or* Y alone, which is called the **marginal probability mass function of** X or Y and defined by

$$f_X(x) = \sum_{y \in D_Y} f(x, y) = P(X = x), \qquad x \in D_X = \{\text{all possible values of } X \text{ in } D\}.$$

$$f_Y(y) = \sum_{x \in D_X} f(x, y) = P(Y = y), \qquad y \in D_Y = \{\text{all possible values of } Y \text{ in } D\}.$$

Definition [independent Random Variables]

The random variables X and Y are **independent** if and only if, for every $x \in D_X$ and $y \in D_Y$,

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

$A \cap B$ Event A Event B

or equivalently,

$$f(x, y) = f_X(x) f_Y(y).$$

otherwise, X and Y are said to be **dependent**.

Example 2

Let the joint *pmf* of X and Y be defined by

$$f(x,y) = \frac{x+y}{21}, \qquad x = 1,2,3, \qquad y = 1,2.$$

Check if *RV X* and *Y* are independent.

Solution:

$$f_X(x) = \sum_{y \in D_Y} f(x,y) = \sum_{y=1}^{2} \frac{x+y}{21} = \frac{2x+3}{21}, \qquad x = 1,2,3.$$

$$f_Y(y) = \sum_{x \in D_X} f(x,y) = \sum_{x=1}^{3} \frac{x+y}{21} = \frac{3y+6}{21}, \qquad y = 1,2.$$

$$f(x,y) = \frac{x+y}{21} \neq \frac{2x+3}{21} \cdot \frac{3y+6}{21} = f_X(x) f_Y(y) \Rightarrow X \text{ and } Y \text{ are dependent.}$$

What's the interpretation of $f_X(x)$ and $f_Y(y)$ and independence?

Consider the *conditional pmf*:

$$f(y|x) = P(Y = y | X = x) = \frac{f(x,y)}{f_X(x)};\ f(x|y) = P(X = x | Y = y) = \frac{f(x,y)}{f_Y(x)}$$

➢ Expectation

Let X_1 and X_2 be discrete *RV* with their joint *pmf* $f(x_1, x_2)$: D → [0,1]. Consider a function $u(x_1, x_2)$ of x_1 *and* x_2. Then:

Expectations of functions of bivariate RVs are computed just as with univariate RVs.

(a) The **mathematical expectation** of $u(X_1, X_2)$, if exists, is given by

$$E\left[u(X_1, X_2)\right] = \sum_{(x_1, x_2)\in D} u(x_1, x_2) f(x_1, x_2).$$

$$E(X) = \sum_{(x,y)\in \bar{S}} xf(x,y) = \sum_{x\in \bar{S}_X} xf_X(x).$$

(b) If $u_i(X_1, X_2) = X_i$ for $i = 1, 2$, then

$$E\left[u_i(X_1, X_2)\right] = E(X_i) = u_i$$

is called the **mean** of X_i for $i = 1, 2$.

(c) If $u_i(X_1, X_2) = (X_i - u_i)^2$ for $i = 1, 2$, then

$$E\left[u_i(X_1, X_2)\right] = E\left[(X_i - u_i)^2\right] = \sigma_i^2 = Var(X_i)$$

is called the **variance** of X_i for $i = 1, 2$.

Example 1 - revisited

Recall that X and Y are discrete *RVs* with joint *pmf* $f(X,Y): \mathrm{D} \to [0,1]$ with $D_X = D_Y = \{1,\ 2,\ 3,\ 4,\ 5,\ 6\}$

$$f(x,y) = \begin{cases} 2/36, & 1 \le x < y \le 6 \\ 1/36, & 1 \le x = y \le 6 \end{cases}$$

Compute $E(X+Y)$:

Solution:

$$E(X+Y) = \sum_{(x,y)\in D} (x+y) f(x,y) = \sum_{1\le x=y\le 6} (x+y)\cdot\frac{1}{36} + \sum_{1\le x<y\le 6} (x+y)\frac{2}{36}$$

$$= \sum_{x=1}^{6} 2x\cdot\frac{1}{36} + \sum_{x=1}^{6}\sum_{y=x+1}^{6} (x+y)\cdot\frac{2}{36} = \frac{252}{36}.$$

Work it by yourself!

Section 4.2 The correlation coefficient

recall that for $u(X,Y)$, its expectation $E\left[u(X,Y)\right]=\sum_{(x,y)\in D} u(x,y)f(x,y).$

Definition [Covariance of X and Y]

Take $u(X,Y)=\left[X-E(X)\right]\left[Y-E(Y)\right]$

$$E\left[(X-E(X))(Y-E(Y))\right]=Cov(X,Y),$$

which is called the covariance of X and Y.

- Motivation: To study the relation between X and Y.

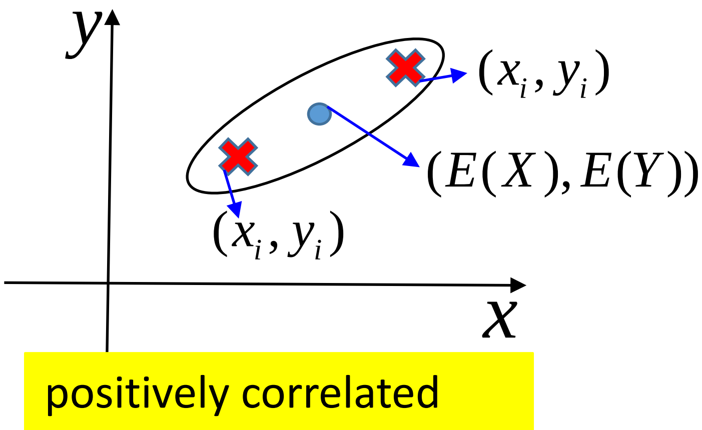
- $Cov(X,Y)=E(XY)-E(X)E(Y)$ → Verify it by yourself!

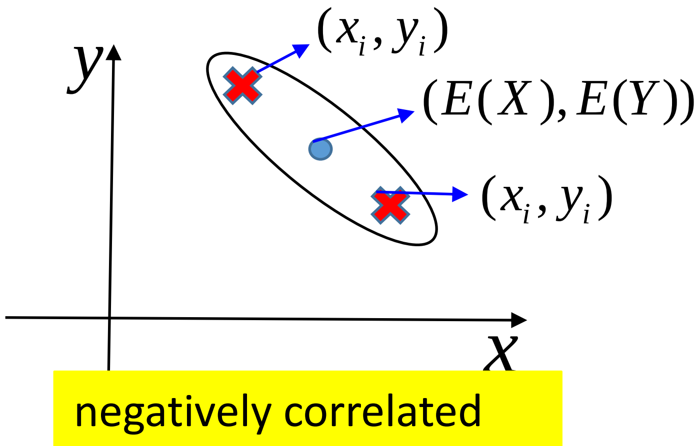
When $Cov(X,Y)=0$, we say X *and* Y are ***uncorrelated***.

- Interpretation: Roughly speaking, a positive or negative covariance indicates 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.

Example 1: Demonstration of positively correlated and negatively correlated RVs

Assume that X and Y are uniformly distributed over the ellipses.





Independence of X and Y could imply the **uncorrelation** of X and Y.

Consider the case that X and Y are independent:

$$E(XY) = \sum_{(x,y)\in D} xyf(x,y) = \sum_{x\in D_X} \sum_{y\in D_Y} xyf_X(x) f_Y(y)$$

$$= \sum_{x\in D_X} xf_X(x) \left[\sum_{y\in D_Y} yf_Y(y) \right] = E(X)E(Y).$$

$f(x, y) = f_X(x) f_Y(y)$
$\Rightarrow D = D_X D_Y$

Therefore, $cov(X,Y) = E(XY) - E(X)E(Y) = 0.$

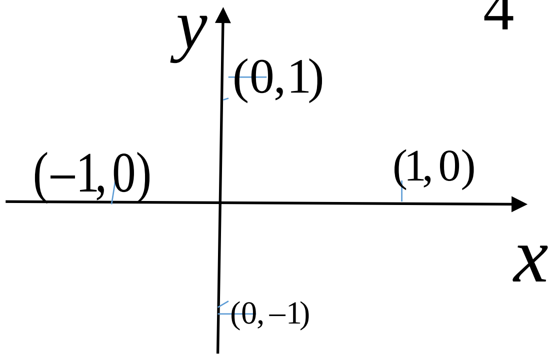
Independent of 2 RVs $\Rightarrow$ uncorrelation of 2 RVs.

However, the converse is not true, that is to say, there exists X and Y which are ***uncorrelated*** but ***not independent***.

Example 2 (**uncorrelation doesn't imply independence**)

Let X and Y be RVs that take values (1, 0), (0, 1), (-1, 0), (0, -1)

and with probability $\frac{1}{4}$, as shown in the figure below.



Q1 : what are the **marginal pmf** of X and Y?

Q2 : what is $Cov(X,Y)$?

Q3 : Are X and Y *independent*?

Solution : To find marginal pmf of X and Y, $D_X = D_Y = \{-1, 0, -1\}$.

$$f_X(x) = \begin{cases} 1/4, & x = 1 \\ 1/2, & x = 0 \\ 1/4, & x = -1 \end{cases}, \qquad f_Y(y) = \begin{cases} 1/4, & y = 1 \\ 1/2, & y = 0 \\ 1/4, & y = -1 \end{cases}$$

$$Cov(X,Y) = E(XY) - E(X)E(Y) = 0 - 0 \cdot 0 = 0.$$

$$f_X(0)\, f_Y(1) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8} \neq f(0,1) = \frac{1}{4} \Rightarrow X \text{ and } Y \text{ are } \textit{not independent}!$$

Definition [correlation coefficients]

The correlation coefficients of X and Y that have nonzero variance is defined as

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

- It is a normalized version of $Cov(X,Y)$ and in fact $-1 \le \rho \le 1$
- Interpretation: $\rho > 0$ $(or\ \rho < 0)$ indicate the values of $X - E(X)$ and $Y - E(Y)$ 'tend' to have the same(or opposite, respectively) sign.
- $\rho > 0$ $(or\ \rho < 0)$ have the same interpretation as $Cov(X,Y) > 0$ (or $Cov(X,Y) < 0$)
- 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

$$Y - E(Y) = c\left[X - E(X)\right]$$

Example 3

Consider *n* **independent** tosses of a coin with probability of a head equal to *p*. Let *X* and *Y* be the number if heads and of tails, respectively. Calculate the correlation coefficient of ***X*** and ***Y*.**

Solution:

$$X+Y=n \Rightarrow E(X)+E(Y)=n \Rightarrow X-E(X)=-\left[Y-E(Y)\right]$$

$$Cov(X,Y)=E\left[(X-E(X))(Y-E(Y))\right]=-E\left[(Y-E(Y))^2\right]=-Var(Y)$$

$$Var(X)=E\left[(X-E(X))^2\right]=E\left[(Y-E(Y))^2\right]=Var(Y)$$

$$\Rightarrow \rho(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}=\frac{-Var(Y)}{\sqrt{Var(Y)}\sqrt{Var(Y)}}=-1.$$

Chapter 4 Bivariate Distributions

Section 4.3 CONDITIONAL DISTRIBUTIONS

➢ Motivation

- Let X and Y have the joint pmf $f(x, y)$: D → [0,1].
- The marginal *pmf* of X and Y are

$$f_X(x)\colon D_X \to [0,1] \text{ and } f_Y(y)\colon D_Y \to [0,1].$$

- By definition,

$$f(x, y) = P(X = x, Y = y) \triangleq P(\{X = x, Y = y\}).$$

$$f_X(x) = P(X = x) \triangleq P(\{X = x\}) = \sum_{y \in D_Y} f(x, y).$$

$$f_Y(y) = P(Y = y) \triangleq P(\{Y = y\}) = \sum_{x \in D_X} f(x, y).$$

- Let $A = \{X = x\},\ B = \{Y = y\}$,
  $A \cap B = \{X = x\} \cap \{Y = y\} \triangleq \{X = x, Y = y\}$.

Recall the conditional probability of event A given event B is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{f(x, y)}{f_Y(y)} \text{ (under the assumption } f_Y(y) > 0).$$

Definition [conditional probability mass function]

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

$$g(x|y)=\frac{f(x,y)}{f_Y(y)}, \qquad \text{provided that } f_Y(y)>0$$

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

$$h(y|x)=\frac{f(x,y)}{f_X(x)}, \qquad \text{provided that } f_X(x)>0$$

Example 1: Let the joint *pmf* of X and Y be defined by

$$f(x,y)=\frac{x+y}{21}, \qquad x=1,2,3, \qquad y=1,2.$$

We have shown

$$f_X(x)=\sum_{y\in D_Y} f(x,y)=\sum_{y=1}^{2}\frac{x+y}{21}=\frac{2x+3}{21}, \qquad x=1,2,3.$$

$$f_Y(y)=\sum_{x\in D_X} f(x,y)=\sum_{x=1}^{3}\frac{x+y}{21}=\frac{y+2}{7}, \qquad y=1,2.$$

Then the conditional *pmf* of X given $Y=y$ is

$$g(x|y)=\frac{f(x,y)}{f_Y(y)}=\left(\frac{x+y}{21}\right)\Big/\left(\frac{y+2}{7}\right)=\frac{x+y}{3(y+2)}, \qquad x=1,2,3 \qquad y=1,2.$$

and the conditional *pmf* of Y given $X=x$ is

$$h(y|x)=\frac{f(x,y)}{f_X(x)}=\left(\frac{x+y}{21}\right)\Big/\left(\frac{2x+3}{21}\right)=\frac{x+y}{2x+3}, \qquad x=1,2,3 \qquad y=1,2.$$

➢ Conditional *pmf* is a well-defined *pmf*

① $h(y|x) \geq 0.$

② $\sum_{y \in D_Y} h(y|x) = \sum_{y \in D_Y} \frac{f(x,y)}{f_X(x)} = \frac{\sum_{y \in D_Y} f(x,y)}{f_X(x)} = \frac{f_X(x)}{f_X(x)} = 1.$

$f(x,y) = 0$ if $(x,y) \notin D.$

➢ Conditional mean and conditional variance

Let $u(Y)$ be a function of Y. Then the **conditional expectation** of $u(Y)$ is given by

$$E(u(Y)|X=x) = \sum_{y \in D_Y} u(y) h(y|x).$$

When $u(Y) = Y$,

$$E(Y|X=x) = \sum_{y \in D_Y} y h(y|x).$$

Conditional mean

When $u(Y) = \left[Y - E(Y|X=x)\right]^2$,

$$Var(Y|X=x) \triangleq E\left\{\left[Y - E(Y|X=x)\right]^2 \middle| X=x\right\} = \sum_{y \in D_Y} \left[y - E(Y|X=x)\right]^2 h(y|x).$$

Conditional variance

Example 1 (c.n.t.)

$$E(Y|X=3) = \sum_{y \in D_Y} y h(y|3) = \sum_{y=1}^{2} y \cdot \frac{y+3}{9} = \frac{14}{9}.$$

$$Var(Y|X=3) = \sum_{y \in D_Y} \left[y - E(Y|X=3)\right]^2 h(y|3) = \sum_{y=1}^{2} \left(y - \frac{14}{9}\right)^2 \cdot \frac{y+3}{9} = \frac{20}{81}.$$

Section 4.4 Bivariate Distribution of continuous type

- Idea: (bivariate) ***discrete*** RV → (bivariate) ***continuous*** RV

Definition [joint probability density function (joint *pdf*)]

Let X and Y be two **continuous** *RV*s. The function $f(x, y)$: D → [0, +∞) is called the **joint probability density function** (**joint *pdf***) of X *if*:

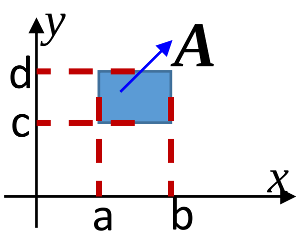
Motivation: The outcome is a tuple of 2 scalars whose range are intervals or union of intervals

① $f(x, y) \geq 0; \quad (x, y) \in D.$

② $\iint_D f(x, y)dxdy = 1;$

③ $P[(X, Y) \in A] \triangleq P(\{(x, y) \in A\}) = \iint_A f(x, y)dxdy, \quad A \subseteq D.$

Remark :



- Very often, we extend the definition domain of $f(x, y)$ from D to $R \times R$ by letting $f(x, y) = 0$ for $(x, y) \notin D$ and thus $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x, y)dxdy = 1.$
- In this course, we only consider a special space A in the ③ of the definiton: A is ***rectangular*** with its line segments parallel to the coordinate axis. In this case, $A = \{(x, y) | a \leq x \leq b, \ c \leq y \leq d\}$. Then the double integral becomes

$$P((x, y) \in A) = \int_a^b \int_c^d f(x, y)dydx.$$

Remark3 : Joint pdf can be seen as an extension of joint pmf by extending the 'summation' to 'integral'.

- mass → density
- pmf → pdf
- joint pmf → joint pdf
- marginal pmf → marginal pdf
- conditional pmf → conditional pdf

summation → integral

Mean

Variance

Covariance

Correlation

Definition [Marginal *pdf*]

The **marginal probability density function of** X or Y is defined by

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy : D_X \to [0,+\infty)$$
$$x \in D_X = \{\text{all possible values of } x \text{ in } D\}.$$
$$f_X(x) : R \to [0,+\infty) \text{ by letting } f_X(x) = 0 \text{ for } x \notin D_X$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx : D_Y \to [0,+\infty)$$
$$y \in D_Y = \{\text{all possible values of } y \text{ in } D\}.$$
$$f_Y(y) : R \to [0,+\infty) \text{ by letting } f_Y(y) = 0 \text{ for } y \notin D_Y$$

Definition [Mathematical expectation]

Let $u(X,Y)$ be a function of X and Y whose marginal pdf is given by $f(x,y)$. Thus the mathematical expectation of $u(X,Y)$ is defined by

$$E[u(X,Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} u(x,y) f(x,y)dxdy$$

➢ When $u(X,Y)=X$,

$$E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xf(x,y)\,dx\,dy=\int_{-\infty}^{+\infty} xf_X(x)\,dx.$$

➢ When $u(X,Y)=(X-E(X))^2$,

$$Var(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(X-E(X))^2 f(x,y)\,dx\,dy=\int_{-\infty}^{+\infty}(X-E(X))^2 f_X(x)\,dx.$$

Example 1

Let X and Y have the joint pdf $f(x,y)=\frac{4}{3}(1-xy)$ with $0\le x\le 1,\ 0\le y\le 1.$

Compute $f_X(x)$, $f_Y(y)$, $E(X)$ and $Var(X)$.

Solution:

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\int_0^1 \frac{4}{3}(1-xy)dy=\frac{4}{3}-\frac{4}{3}(x)(\frac{1}{2}y^2)\Big|_0^1=\frac{4}{3}(1-\frac{1}{2}x)$$

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\int_0^1 \frac{4}{3}(1-xy)dx=\frac{4}{3}(1-\frac{1}{2}y) \quad \leftarrow \text{Due to the symmetry.}$$

$$E(X)=\int_{-\infty}^{+\infty} xf_X(x)dx=\int_0^1 x\frac{4}{3}(1-\frac{1}{2}x)dx=\frac{4}{3}\left[\frac{1}{2}x^2\Big|_0^1-\frac{1}{6}x^3\Big|_0^1\right]=\frac{4}{9}$$

$$Var(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2 f_X(x)dx=\int_0^1 (x-\frac{4}{9})^2\frac{4}{3}(1-\frac{1}{2}x)dx=\frac{13}{162}.$$

You should verify the details by yourself!

Quiz

Let X and Y have the joint pdf $f(x,y)=\frac{3}{2}x^2(1-|y|)$ with $-1<x<1,\ -1<y<1$.

$A=\{(x,y)|0<x<1, 0<y<x\}$. Compute $E(X)$ and $P(A)$.

Solution:

$$f_X(x)=\int_{-1}^{1}\frac{3}{2}x^2(1-|y|)dy=\int_{0}^{1}\frac{3}{2}x^2(1-y)dy+\int_{-1}^{0}\frac{3}{2}x^2(1+y)dy$$

$$=\frac{3}{2}x^2\left[y-\frac{1}{2}y^2\right]_0^1+\frac{3}{2}x^2\left[y+\frac{1}{2}y^2\right]_{-1}^0=\frac{3}{2}x^2\times\frac{1}{2}+\frac{3}{2}x^2\times\frac{1}{2}=\frac{3}{2}x^2.$$

$$E(X)=\int_{-1}^{1}xf_X(x)dx=\int_{-1}^{1}\frac{3}{2}x^3dx=\left[\frac{3}{8}x^4\right]_{-1}^{1}=0.$$

We have two ways to compute $P(A)$:

$$P(A)=\iint_A f(x,y)dxdy=\int_0^1\int_0^x\frac{3}{2}x^2(1-|y|)dydx=\int_0^1\int_0^x\frac{3}{2}x^2(1-y)dydx=\int_0^1\left[\frac{3}{2}x^2(y-\frac{1}{2}y^2)\right]_0^x dx$$

$$=\int_0^1(\frac{3}{2}x^3-\frac{3}{4}x^4)dx=\left[\frac{3}{8}x^4-\frac{3}{20}x^5\right]_0^1=\frac{9}{40}.$$

$$P(A)=\iint_A f(x,y)dydx=\int_0^1\int_y^1\frac{3}{2}x^2(1-|y|)dxdy=\int_0^1\left[\frac{x^3}{2}(1-|y|)\right]_y^1 dy=\int_0^1(\frac{1}{2}-\frac{y^3}{2})(1-|y|)dy$$

$$=\int_0^1\left[\frac{y^4}{2}-\frac{y^3}{2}-\frac{y}{2}+\frac{1}{2}\right]dy=\left[\frac{y^5}{10}-\frac{y^4}{8}-\frac{y^2}{4}+\frac{1}{2}y\right]_0^1=\frac{1}{10}-\frac{1}{8}-\frac{1}{4}+\frac{1}{2}=\frac{9}{40}.$$

Definition [independent Continuous Variables]

Two continuous variables X and Y are **independent** if and only if,

$$f(x,y) = f_X(x)\, f_Y(y), \qquad x \in D_X,\, y \in D_Y$$

Otherwise, X and Y are said to be **dependent**.

Example 1 (Revisited)

Since $f(x,y) = \frac{4}{3}(1-xy) \neq \left[\frac{4}{3}(1-\frac{1}{2}x)\right]\left[\frac{4}{3}(1-\frac{1}{2}y)\right] = f_X(x)\, f_Y(y)$, X and Y are *dependent*.

Definition [Covariance and correlation coefficient]

The **covariance** of X and Y is given by

$$Cov(X,Y) = E\left[(X - E(X))(Y - E(Y))\right] = E(XY) - E(X)E(Y),$$

$$\text{where } E(XY) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x,y)\,dxdy$$

The correlation coefficients is defined as

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

Definition 4.4-6 [conditional probability density function]
Let X and Y have the joint pdf $f(x, y)$ and marginal pdfs are $f_X(x)$ and $f_Y(y)$.
Then the **conditional pdf**, **mean**, and **variance** of Y, given that $X=x$, are

$$h(y|x) = \frac{f(x,y)}{f_X(x)} \text{ for } f_X(x) > 0,$$

$$E(Y|X = x) = \int_{-\infty}^{+\infty} yh(y|x)dy,$$

$$Var(Y|X = x) = E([Y - E(Y|X = x)]^2 |X = x) = \int_{-\infty}^{+\infty} (y - E(Y|X = x))^2 h(y|x)dy$$

$$= E(Y^2|X = x) - [E(Y|X = x)]^2$$

Example 2

Let X and Y be continuous RVs that have

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

Question:

(a) Sketch the support of X and Y.

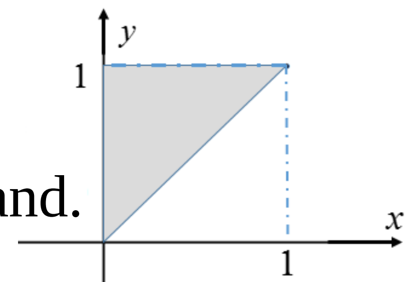
(b) Compute the marginal pmfs $f_X(x)$ and $f_Y(y)$.

(c) Compute the conditional pdf, conditional mean, conditional variance of Y, given $X = x$.

(d) Compute $P(\frac{3}{4} \le Y \le \frac{7}{8} \Big| X = \frac{1}{4})$.

Example 2 (c.n.t.)

Solution :



(a) The graph for the support of X and Y is listed righthand.

(b) $f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\int_x^1 f(x,y)dy=2(1-x) \qquad 0\le x\le 1.$

$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\int_0^y f(x,y)dy=2y \qquad 0\le y\le 1.$

(c) $h(y|x)=\dfrac{f(x,y)}{f_X(x)}=\dfrac{2}{2(1-x)}=\dfrac{1}{1-x}, \quad 0\le x\le y\le 1.$

$$E(Y|X=x)=\int_{-\infty}^{+\infty} yh(y|x)dy=\int_x^1 y\frac{1}{1-x}dy=\frac{1}{1-x}\left[\frac{1}{2}y^2\right]_x^1=\frac{1}{2}(x+1).$$

$$Var(Y|X=x)=\int_{-\infty}^{+\infty}\left[y-E(Y|X=x)\right]^2 h(y|x)dy$$

$$=\int_x^1\left[y-\frac{1}{2}(x+1)\right]^2\frac{1}{1-x}dy=\frac{1}{1-x}\left[y-\frac{1}{2}(1+x)\right]^3\Big|_x^1=\frac{(1-x)^2}{12}.$$

(d) $P(\frac{3}{4}\le Y\le\frac{7}{8}\Big|X=\frac{1}{4})=\int_{3/4}^{7/8} h(y\Big|\frac{1}{4})dy=\int_{3/4}^{7/8}\frac{1}{1-1/4}dy=\frac{1}{8}\times\frac{4}{3}=\frac{1}{6}.$

Section 4.5 Bivariate Normal Distribution

Definition [Bivariate Normal]

Let X and Y be two continuous *RVs* and have the joint pdf

$$f(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-p^2}}\exp\left[-\frac{1}{2}q(x,y)\right]$$

where $q(x,y)=\frac{1}{1-\rho^2}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2-2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\geq 0$

$\mu_X=E(X),\mu_Y=E(Y),\sigma_X=\sqrt{Var(X)},\sigma_Y=\sqrt{Var(Y)},\rho$ is the **correlation coefficient**.

Then X and Y are said to be **bivariate normal distributed.**

➢ Property:

Given that $Y=y$ is normal distribution, The probability distribution of X with **mean** $\mu_X+\frac{\sigma_X}{\sigma_Y}\rho(y-\mu_Y)$ and **variance** $(1-\rho^2)\sigma_X^2$ is given by

$$X\,|\,Y=y\sim N(\mu_X+\frac{\sigma_X}{\sigma_Y}\rho(y-\mu_Y),(1-\rho^2)\sigma_X^2).$$

Similarly, $$Y\,|\,X=x\sim N(\mu_Y+\frac{\sigma_Y}{\sigma_X}\rho(x-\mu_X),(1-\rho^2)\sigma_Y^2).$$

Example 1

Observe a group of college students, Let X and Y denote the grade points in high school and the first year in college have a bivariate normal distribution with parameters

$\mu_X = 2.9 \qquad \mu_Y = 2.4 \qquad \sigma_X = 0.4 \qquad \sigma_Y = 0.5 \qquad \rho = 0.8$

Compute $P(2.1 < Y < 3.3)$ and $P(2.1 < Y < 3.3 \mid X = 3.2)$

Solution:

$$P(2.1 < Y < 3.3) = P\left(\frac{2.1-2.4}{0.5} < \frac{Y-2.4}{0.5} < \frac{3.3-2.4}{0.5}\right) = \Phi(1.8) - \Phi(-0.6) = 0.69$$

Note that $Y \mid X = x \sim N\left(\mu_Y + \frac{\sigma_Y}{\sigma_X}\rho(x-\mu_X), (1-\rho^2)\sigma_Y^2\right)$,

when $X = 3.2$, $Y \mid X = 3.2 \sim N(2.7, 0.09)$,

$$P(2.1 < Y < 3.3 \mid X = 3.2) = P\left(\frac{2.1-2.7}{\sqrt{0.09}} < \frac{Y-2.7}{\sqrt{0.09}} < \frac{3.3-2.7}{\sqrt{0.09}}\right)$$

$$= \Phi(2) - \Phi(-2) = 0.95.$$

Theorem [Bivariate Normal: Uncorrelation Implies Independence]
If X and Y have a bivariate normal distribution with correlation coefficient ρ, then X and Y are independent if and only if $\rho = 0$.