50.034 - Introduction to Probability and Statistics

Week 4 – Lecture 7

January–May Term, 2019

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Lecture 7 Assigned Readings: 3.4, 3.5

Outline of Lecture

- Joint distribution
- Joint pmf and joint pdf
- Marginal pmf and marginal pdf
- Independence of R.V.'s
- Joint cumulative distribution function

Question

Roll a die twice. Let T_1, T_2 be R.V.'s representing the outcomes of the two rolls. Is there any relationship between T_1 and T_2?

- If the die is fair and rolled randomly, there is no relationship between T_1 and T_2, i.e. T_1, T_2 are independent R.V.'s.
- **Recall:** T_1 and T_2 are independent if for all sets C_1, C_2 of real numbers, the events $\{T_1 \in C_1\}$, $\{T_2 \in C_2\}$ are independent.

Now, let $X = \min\{T_1, T_2\}$ and $Y = \max\{T_1, T_2\}$.

Is there any relationship between the two new R.V.'s X and Y?

- Yes. For example, if $X = 2$, then it must be that $Y \geq 2$.
- X and Y are not independent, i.e. X and Y are dependent.

To better understand how two dependent R.V.'s depend on each other, we need to consider their distributions **jointly**.

Joint distributions

Recall: The probability distribution of *any* R.V. X is the collection of all probabilities of the form $\Pr(X \in C)$, for all sets $C \subseteq \mathbb{R}$.

- **Interpretation:** For any set $C \subseteq \mathbb{R}$, this distribution gives the probability $\Pr(X \in C)$ of how likely X takes on values in C.
- "distribution of X" = "probability distribution of X".

Let X and Y be *any* R.V.'s defined on the sample space Ω. The joint distribution of X and Y is the collection of all probabilities of the form $\Pr\left((X, Y) \in C\right)$, for all sets $C \subseteq \mathbb{R}^2$.

- C contains pairs of real numbers.
- $\{(X, Y) \in C\}$ is the set $\{\omega \in \Omega : (X(\omega), Y(\omega)) \in C\}$ of all outcomes whose X-value and Y-value form a pair in C.
- So $\{(X, Y) \in C\}$ is an event.
- $\Pr\left((X, Y) \in C\right)$ is the probability of the event $\{(X, Y) \in C\}$.

More on joint distributions

Interpretation: For any set $C \subseteq \mathbb{R}^2$, the joint distribution gives the probability $\Pr\big((X, Y) \in C\big)$ of how likely (X, Y) takes on a pair of values that appears in C.

Remarks on notation:

- "Joint distribution" = "joint probability distribution".
- We write $\Pr(X = x, Y = y)$ to mean $\Pr\big((X, Y) \in \{(x, y)\}\big)$.

There are other ways to represent the same information given by the joint distribution of two R.V.'s:

- joint probability mass function (only for discrete R.V.'s)
- joint probability density function (only for continuous R.V.'s)
- joint cumulative distribution function (for *any* R.V.'s)

Example 1

Consider the roll outcome of a fair die. Let

$$X = \begin{cases} 1, & \text{if the outcome is even;} \\ 0, & \text{otherwise;} \end{cases}$$

$$Y = \begin{cases} 1, & \text{if the outcome is prime;} \\ 0, & \text{otherwise.} \end{cases}$$

What is $\Pr(X = 1, Y = 1)$?

Answer: $\frac{1}{6}$. Only one outcome (the outcome 2), among the six possible outcomes, satisfies both $X = 1$ and $Y = 1$.

Joint probability mass function

Let X and Y be two **discrete** R.V.'s.
The joint probability mass function (joint pmf) $p(x,y)$ is defined for each pair of numbers (x,y) by

$$p(x,y) = \Pr(X = x \text{ and } Y = y) = \Pr\left((X,Y) = (x,y)\right).$$

A legitimate joint pmf must satisfy

$$p(x,y) \geq 0 \text{ and } \sum_x \sum_y p(x,y) = 1.$$

Given any subset $A \subseteq \mathbb{R} \times \mathbb{R}$, the probability $\Pr\left((X,Y) \in A\right)$ is obtained by summing the joint pmf over all pairs in A:

$$\Pr\left((X,Y) \in A\right) = \sum_{(x,y)\in A} p(x,y)$$

Marginal probability mass function

Let X and Y be two **discrete** R.V.'s with joint pmf $p(x, y)$, and suppose their sets of possible values are D_X and D_Y respectively.

The marginal probability mass function (marginal pmf) of X, denoted by $p_X(x)$, is given by

$$p_X(x) = \sum_{y \in D_Y} p(x, y)$$

for each possible value $x \in D_X$.

Similarly, the marginal pmf of Y is

$$p_Y(x) = \sum_{x \in D_X} p(x, y)$$

for each possible value $y \in D_Y$.

Example 2

Let X and Y be discrete R.V.'s, whose joint pmf is given by:

	$Y = 0$	1	2	3
$\mathrm{X} = 0$	$1/8$	$2/8$	$1/8$	0
1	0	$1/8$	$2/8$	$1/8$

Then, the marginal pmf of X is:

X	0	1
$p_X(x)$	$1/2$	$1/2$

And, the marginal pmf of Y is:

Y	0	1	2	3
$p_Y(y)$	$1/8$	$3/8$	$3/8$	$1/8$

- The marginal pmf's can be found by summing along the **margins** of the joint pmf table.
- This is the origin of why the pmf is called "marginal".

Remarks on marginal probability mass functions

Let X and Y be **discrete** R.V.'s.

The marginal pmf of X is exactly the pmf of X. Similarly, the marginal pmf of Y is exactly the pmf of Y.

- The word "marginal" indicates that the pmf is obtained from a joint probability distribution (by summing the joint pmf table along the margin for X).

A marginal pmf must satisfy the same conditions as a pmf:

- $p_X(x) \geq 0$ and $p_Y(y) \geq 0$ for all x, y.
  (All probabilities are non-negative.)
- $\sum_x p_X(x) = 1$ and $\sum_y p_Y(y) = 1$.
  (The sum of probabilities must be 1.)

From a joint pmf, we can obtain the marginal pmf of each R.V.; however, the converse is not always true.

Joint probability density function

Let X and Y be two continuous R.V.'s.
A joint probability density function (joint pdf) of X and Y is a function $f(x,y)$ satisfying the following:

- $f(x,y)$ is a non-negative function, i.e. $f(x,y) \geq 0$ for all x, y
- For any set $A \subseteq \mathbb{R}^2$, the probability of event $\{(X,Y) \in A\}$ is

$$\Pr\left((X,Y) \in A\right) = \iint_A f(x,y)\,dx\,dy.$$

(Here, A is a set consisting of pairs of real numbers.)
Note: $\Pr\left((X,Y) \in \mathbb{R}^2\right) = 1$, so this joint pdf must satisfy

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1.$$

We can think of $\iint_A f(x,y)\,dx\,dy$ as the volume of the region under the graph of $f(x,y)$. Here, the graph of $f(x,y)$ forms a surface over the xy-plane.

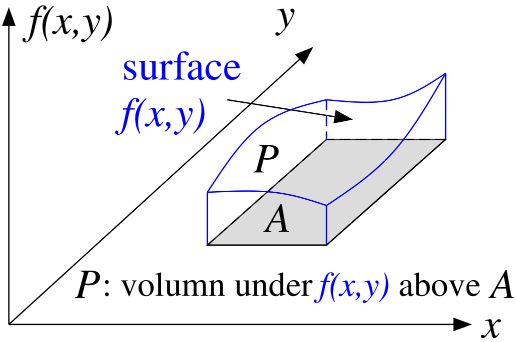
Joint probability density function

In particular, if A is the two-dimensional rectangular region

$$A = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, c \leq y \leq d\},$$

then

$$\Pr((X, Y) \in A) = \int_c^d \int_a^b f(x, y)\, dx\, dy.$$



Marginal probability density function

Let X and Y be two **continuous** R.V.'s with joint pdf $f(x, y)$.
The marginal probability density function (marginal pdf) of X and Y, denoted by $f_X(x)$ and $f_Y(y)$ respectively, are

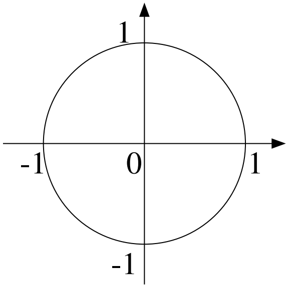
$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy, \text{ for } -\infty < x < \infty;$$
$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx, \text{ for } -\infty < y < \infty.$$

Remarks similar to the marginal pmf case:

- The marginal pdf of X is exactly the pdf of X, and similarly, the marginal pdf of Y is exactly the pdf of Y.
- The word "marginal" indicates that the pdf is obtained from a joint probability distribution.
- A marginal pdf must satisfy the same conditions as a pdf:
  - $f_X(x) \geq 0$, $f_Y(y) \geq 0$ for all x, y. (Density is non-negative.)
  - $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$, $\int_{-\infty}^{\infty} f_Y(y)\, dy = 1$. (Area under curve is 1.)
- From a joint pdf, we can obtain the marginal pdf of each R.V.; however, the converse is not always true.

Example 3

A point is chosen randomly from a disk of radius 1. Let X and Y denote the x-coordinate and y-coordinate respectively of the point.
The joint pdf of X and Y is



$$f(x,y) = \begin{cases} c, & \text{if } x^2 + y^2 \le 1; \\ 0, & \text{otherwise;} \end{cases}$$

for some constant c.

(1) What is the value of c?

(2) What is the marginal pdf of Y?

Example 3

(1): Since $f(x,y)$ is a legitimate pdf, it must satisfy

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

or equivalently,

$$\int_{-1}^{1}\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} c\,dy\,dx = \pi c = 1.$$

Note: If $f(x,y)$ is a constant not related to x or y, then the result of integration is the area of the region times the constant.

Therefore, $c = \frac{1}{\pi}$.

Example 3

(2): We check that

$$\int_{-\infty}^{\infty} f(x,y)\,dx = \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{1}{\pi}\,dx$$
$$= \frac{2}{\pi}\sqrt{1-y^2}$$

if $-1 \le y \le 1$, and $\int_{-\infty}^{\infty} f(x,y)\,dx = 0$ if $y < -1$ or $y > 1$.

Therefore, the marginal pdf of Y is

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \begin{cases} \frac{2}{\pi}\sqrt{1-y^2}, & \text{if } -1 \le y \le 1; \\ 0, & \text{otherwise.} \end{cases}$$

Recall: Independence

What does the independence of two events A and B mean?

It means the occurrence of A has no bearing on B and vice versa.

If A and B are independent, then

$$\Pr(A|B) = \Pr(A) \text{ or}$$

$$\Pr(A \cap B) = \Pr(A)\Pr(B)$$

Independence of R.V.'s

Recall: Two R.V.'s X and Y are called independent if for every two sets C, C' of real numbers,

$$\Pr(X \in C, Y \in C') = \Pr(X \in C)\Pr(Y \in C').$$

- The comma in "$\Pr(X \in C, Y \in C')$" means "and".
- $\Pr(X \in C, Y \in C') = \Pr\left(\{X \in C\} \cap \{Y \in C'\}\right)$.

There are several useful equivalent conditions for independence.

p(x,y) : P(X n Y)

Theorem:

- Two discrete R.V.'s X and Y are independent if and only if their joint pmf $p(x, y)$ is the **product of the marginal pmf's**.
  - i.e. $p(x, y) = p_X(x)p_Y(y)$ for all pairs $(x, y) \in \mathbb{R}^2$.
- Two continuous R.V.'s X and Y are independent if their joint pdf $f(x, y)$ is the **product of the marginal pdf's**.
  - i.e. $f(x, y) = f_X(x)f_Y(y)$ for all pairs $(x, y) \in \mathbb{R}^2$.

Example 4

Let X and Y be discrete R.V.'s whose sets of possible values are $\{1, 2, 3\}$ and $\{5, 6, 7\}$ respectively.

Suppose the joint pmf of X and Y is:

Y \ X	1	2	3
5	0.3	0	0.2
6	0.2	0.1	0.1
7	0	0.1	0

Are X and Y independent?

Example 4

To check whether X and Y are independent, we have to find the marginal pmf's of X and Y.

Y \ X	1	2	3
5	0.3	0	0.2
6	0.2	0.1	0.1
7	0	0.1	0

$\longrightarrow$

$p_X(x)$	1	2	3
	0.5	0.2	0.3
$p_Y(y)$	5	6	7
	0.5	0.4	0.1

Note: $p(1,5) = 0.3$, but $p_X(1)p_Y(5) = 0.25$; hence, X and Y are dependent.

Example 5 (Challenge of the day)

Two counters of a shop, one on each side, serve customers in parallel. Let X and Y be the proportion of a day that each counter is in use (i.e. when at least one customer is being served).
Suppose the joint pdf of X and Y is

$$f(x, y) = \begin{cases} k(x + y^2), & \text{if } 0 \leq x \leq 1, 0 \leq y \leq 1; \\ 0, & \text{otherwise;} \end{cases}$$

for some constant k.

(1) What is the value of k?

(2) What is the probability that neither counter is in use for more than a quarter of a day?

(3) Are X and Y dependent?

Example 5

(1): As a legitimate pdf, $f(x, y)$ must satisfy

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

or equivalently,

$$\begin{aligned}
\int_0^1 \int_0^1 k(x + y^2)\, dx\, dy &= k \cdot \int_0^1 \left[\tfrac{1}{2}x^2 + xy^2\right]_{x=0}^{x=1} dy \\
&= k \cdot \int_0^1 \left(\tfrac{1}{2} + y^2\right) dy \\
&= k \cdot \left[\tfrac{1}{2}y + \tfrac{1}{3}y^3\right]_{y=0}^{y=1} \\
&= k\left(\tfrac{1}{2} + \tfrac{1}{3}\right) \\
&= \tfrac{5}{6}k = 1.
\end{aligned}$$

Therefore, $k = \dfrac{6}{5}$.

Example 5

(2): The probability that neither counter is in use for more than a quarter of a day is:

$$\Pr(0 \leq X \leq 0.25, 0 \leq Y \leq 0.25)$$
$$= \int_0^{\frac{1}{4}} \int_0^{\frac{1}{4}} \tfrac{6}{5}(x + y^2)\, dx\, dy = \frac{6}{5} \int_0^{\frac{1}{4}} \left[\tfrac{1}{2}x^2 + xy^2\right]_{x=0}^{x=\frac{1}{4}} dy$$
$$= \frac{6}{5} \int_0^{\frac{1}{4}} \left(\tfrac{1}{32} + \tfrac{1}{4}y^2\right) dy = \frac{6}{5} \cdot \left[\frac{1}{32}y + \frac{1}{12}y^3\right]_{y=0}^{y=\frac{1}{4}}$$
$$= \frac{7}{640}.$$

Example 5

(3): We check that

$$\int_{-\infty}^{\infty} f(x,y)\,dy = \int_0^1 \frac{6}{5}(x+y^2)\,dy = \frac{6}{5}x + \frac{2}{5};$$

$$\int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^1 \frac{6}{5}(x+y^2)\,dx = \frac{3}{5} + \frac{6}{5}y^2;$$

for $0 \le x \le 1$ and $0 \le y \le 1$ respectively, and that

$$\int_{-\infty}^{\infty} f(x,y)\,dy = 0 \quad \text{and} \quad \int_{-\infty}^{\infty} f(x,y)\,dx = 0$$

for all other values. Therefore, the marginal pdf's of X and Y are:

$$f_X(x) = \begin{cases} \frac{6}{5}x + \frac{2}{5}, & \text{if } 0 \le x \le 1; \\ 0, & \text{otherwise;} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{3}{5} + \frac{6}{5}y^2, & \text{if } 0 \le y \le 1; \\ 0, & \text{otherwise.} \end{cases}$$

Note: $f(x,y) \ne f_X(x)f_Y(y)$, so X and Y are dependent.

Joint cumulative distribution function (cdf)

Recall: The cumulative distribution function (cdf) of *any* R.V. X is the function $F(x) = \Pr(X \leq x)$, for $-\infty < x < \infty$.

- $F(x)$ is the probability that the observed X-value is at most x.

Let X and Y be arbitrary R.V.'s defined on the sample space Ω. The joint cumulative distribution function (joint cdf) of X and Y is the function

$$F(x, y) = \Pr(X \leq x, Y \leq y), \quad \text{for } -\infty < x, y < \infty.$$

- For any real numbers x and y, $\{X \leq x, Y \leq y\}$ denotes the event $\{\omega \in \Omega : X(\omega) \leq x \text{ and } Y(\omega) \leq y\}$.
- Hence, the joint cdf $F(x, y)$ is the probability of the event $\{X \leq x, Y \leq y\}$.

Properties of joint cdf

Let X and Y be arbitrary R.V.'s with joint cdf $F(x, y)$.

- For any fixed $y = y_0$, $F(x, y_0)$ (as a function of x) is **non-decreasing**, i.e.

$$\text{If } x_1 < x_2, \text{ then } F(y_0, x_1) \leq F(y_0, x_2).$$

- Similarly, for any fixed $x = x_0$, $F(x_0, y)$ (as a function of y) is **non-decreasing**, i.e.

$$\text{If } y_1 < y_2, \text{ then } F(y_1, x_0) \leq F(y_2, x_0).$$

- For any real numbers a, b, c, d satisfying $a < b$ and $c < d$,

$$\begin{aligned}&\Pr(a < X \leq b \text{ and } c < Y \leq d)\\ &= F(b, d) - F(a, d) - F(b, c) + F(a, c).\end{aligned}$$

- The limits of $F(x, y)$ at $(\pm\infty, \pm\infty)$:

$$\lim_{\substack{x\to-\infty\\ y\to-\infty}} F(x, y) = 0 \quad \text{and} \quad \lim_{\substack{x\to\infty\\ y\to\infty}} F(x, y) = 1.$$

Joint pdf $f(x, y)$ versus joint cdf $F(x, y)$

Theorem: If X and Y are continuous R.V.'s with joint pdf $f(x, y)$ and joint cdf $F(x, y)$, and if the second-order partial derivative

$$\frac{\partial^2 F(x, y)}{\partial x \partial y}$$

exists at the point $(x, y) = (x_0, y_0)$, then

$$f(x_0, y_0) = \left.\frac{\partial^2 F(x, y)}{\partial x \partial y}\right|_{(x,y)=(x_0,y_0)}.$$

- i.e. we can get $f(x, y)$ from $F(x, y)$ (if $\frac{\partial^2 F}{\partial x \partial y}$ exists).

Theorem: If X and Y are continuous R.V.'s with joint pdf $f(x, y)$ and joint cdf $F(x, y)$, then

$$F(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(u, v)\, du\, dv.$$

- i.e. we can get $F(x, y)$ from $f(x, y)$.

Marginal cumulative distribution function

Definition: If X and Y are **arbitrary** R.V.'s with joint cdf $F(x, y)$, then the marginal cdf of X is

$$F_X(x) = \lim_{y \to \infty} F(x, y),$$

and the marginal cdf of Y is

$$F_Y(y) = \lim_{x \to \infty} F(x, y).$$

- Similar as before, the word "marginal" indicates that the cdf is obtained from a joint probability distribution.
- **Fact:** The marginal cdf of X is precisely the cdf of X.
- **Fact:** Also, the marginal cdf of Y is precisely the cdf of Y.

Example 6

Let X and Y be continuous R.V.'s such that (X, Y) must belong to the rectangle in the xy-plane containing all points (x, y) that satisfy $0 \leq x \leq 2$ and $\leq y \leq 2$. Suppose that the joint cdf of X and Y at every point (x, y) in this rectangle is specified as follows:

$$F(x, y) = \frac{1}{16}xy(x + y).$$

What is the cdf $F_X(x)$ of X? What is the cdf $F_Y(y)$ of Y?

Solution:

$$F_X(x) = \lim_{y\to\infty} F(x, y) = \begin{cases} 0, & \text{if } x < 0; \\ \frac{1}{8}x(x + 2), & \text{if } 0 \leq x \leq 2; \\ 1, & \text{if } x > 2. \end{cases}$$

$$F_Y(y) = \lim_{x\to\infty} F(x, y) = \begin{cases} 0, & \text{if } y < 0; \\ \frac{1}{8}y(y + 2), & \text{if } 0 \leq y \leq 2; \\ 1, & \text{if } y > 2. \end{cases}$$

Summary

- Joint distribution
- Joint pmf and joint pdf
- Marginal pmf and marginal pdf
- Independence of R.V.'s
- Joint cumulative distribution function