Week 8: Joint Distributions

Yilin

DigiPen

JOINT DISTRIBUTIONS

Joint Distributions

Chapter Goal

This chapter is devoted to studying how multiple random variables behave with each other.

- Most processes in life are multivariate.
- Joint distributions serve as a tool to model processes with more than one random variable.
- We will focus most of our attention to **discrete** joint distributions rather than continuous.
- The continuous ones require multivariable calculus in most cases.

Drawing from an urn without replacement

Question

We have four balls in an urn, two marked with 1 and two marked with 0. Draw two without replacement. What are the probabilities of each combination of 0 and 1?

Drawing from an urn without replacement

Question

We have four balls in an urn, two marked with 1 and two marked with 0. Draw two without replacement. What are the probabilities of each combination of 0 and 1?

Let X_1 be the first ball and X_2 be the second (now order matters). Compute

$$
\begin{aligned}
P(X_1 = 0 \text{ and } X_2 = 0) &= P(X_2 = 0|X_1 = 0)P(X_1 = 0) = \tfrac{1}{3} \cdot \tfrac{1}{2} = \tfrac{1}{6} \\
P(X_1 = 1 \text{ and } X_2 = 0) &= P(X_2 = 0|X_1 = 1)P(X_1 = 1) = \tfrac{2}{3} \cdot \tfrac{1}{2} = \tfrac{1}{3} \\
P(X_1 = 0 \text{ and } X_2 = 1) &= P(X_2 = 1|X_1 = 0)P(X_1 = 0) = \tfrac{2}{3} \cdot \tfrac{1}{2} = \tfrac{1}{3} \\
P(X_1 = 1 \text{ and } X_2 = 1) &= P(X_2 = 1|X_1 = 1)P(X_1 = 1) = \tfrac{1}{3} \cdot \tfrac{1}{2} = \tfrac{1}{6}.
\end{aligned}
$$

Displaying the Joint Probabilities

We can write this **joint distribution** more neatly in a table:

Joint Prob.	$X_2 = 0$	$X_2 = 1$
$X_1 = 0$	1/6	1/3
$X_1 = 1$	1/3	1/6

We can learn a lot from this information.

- We can compute individual or **marginal** probabilities like

$$P(X_2 = 0) = P(X_1 = 0, X_2 = 0) + P(X_1 = 1, X_2 = 0) = \frac{1}{6} + \frac{1}{3} = \frac{1}{2}.$$

- We see that X_1 and X_2 are **dependent** because, e.g.,

$$P(X_1 = 0, X_2 = 0) = \frac{1}{6}, \text{ but } P(X_1 = 0)P(X_2 = 0) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

Drawing from an urn with replacement

Replacement

Now replace the first ball before drawing the second. Then X_1 and X_2 would be **independent**, since the selection of the first would not affect the selection of the second.

For example,

$$P(X_2 = 1|X_1 = 0) = P(X_2 = 1).$$

From this we would compute that

$$P(X_1 = i, X_2 = j) = P(X_1 = i)P(X_2 = j) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

Thus

Joint Prob.	$X_2 = 0$	$X_2 = 1$
$X_1 = 0$	1/4	1/4
$X_1 = 1$	1/4	1/4

Definition

Definition

Let X and Y be discrete r.v.s, where X takes values in $x_1, \ldots, x_n$ and Y takes values in $y_1, \ldots, y_m$. Then the **joint distribution of** X **and** Y is given by the **joint probabilities**:

$$p_{ij} := P(X = x_i \text{ and } Y = y_j) = P(X = x_i, Y = y_j),$$

$$1 \leq i \leq n, \; 1 \leq j \leq m.$$

For two r.v.s, this is easily displayed in a table:

Joint Prob.	$Y = y_1$	$y = y_2$	$\cdots$	$y = y_m$
$X = x_1$	p_{11}	p_{12}	$\cdots$	p_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$X = n$	p_{n1}	p_{n2}	$\cdots$	p_{nm}

A Joint Game

Flip a coin and roll a six-sided die.

- Let X be 1 if heads, 0 if tails.
- Let Y be the score of the die roll, 1 through 6.
- Assume the flip and the roll are independent.

We can compute the joint probabilities. For $i = 0, 1$, and $j = 1, \ldots, 6$,

$$P(X = i, Y = j) = \overbrace{P(X = i)P(Y = j)}^{\text{Independence}} = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12},$$

We can display the joint distribution

Joint Prob.	$Y = 1$	$Y = 2$	$Y = 3$	$Y = 4$	$Y = 5$	$Y = 6$
$X = 0$	1/12	1/12	1/12	1/12	1/12	1/12
$X = 1$	1/12	1/12	1/12	1/12	1/12	1/12

MARGINAL DISTRIBUTIONS

Marginal Distributions for an urn without replacement

Question

Find the distribution of X_2, that is the **marginal distribution** of X_2.

Solution: Since all the possibilities for X_1 are only 0 or 1, we get

$$P(X_2 = 0) = \overbrace{P(X_1 = 0, X_2 = 0) + P(X_1 = 1, X_2 = 0)}^{X_1 \text{ is either 0 or 1}} = \frac{1}{2}$$
$$P(X_2 = 1) = P(X_1 = 0, X_2 = 1) + P(X_1 = 1, X_2 = 1) = \frac{1}{2}.$$

In a table, the marginal distribution for X_2 is displayed as

$X_2 =$	0	1
Probability	1/2	1/2

This tells us that the probability of the second ball being 0 or 1 is the same!

A Rigged Game

Suppose an enterprising thief sets up a rigged game.

- The mark flips a biased coin and selects a random card from the three: A♠ 2♦ 3♣ .
- Let X be the outcome of the coin flip 1 for heads, 0 for tails
- Let Y represent the card, 1 for A♠ , 2 for 2♦ and 3 for 3♣ .
- Suppose we observe the following joint distribution for X and Y:

Joint Prob.	$Y = 1$	$Y = 2$	$Y = 3$
$X = 0$	0.02	0.31	0.37
$X = 1$	0.03	0.11	0.16

Question

How is X distributed? How is Y distributed?

Marginals for the Rigged Game

Answer

If we know how exactly how Y behaves when $X = i$, we can determine $P(X = i)$:

$$\begin{aligned} P(X=0) &= P(X=0, Y=1) + P(X=0, Y=2) + P(X=0, Y=3) \\ &= 0.02 + 0.31 + 0.37 \\ &= 0.7, \end{aligned}$$

and $P(X = 1) = 1 - P(X = 0) = 0.3$. We can do the same with Y,

$$P(Y=1) = P(X=0, Y=1) + P(X=1, Y=1) = 0.02 + 0.03 = 0.05$$
$$P(Y=2) = P(X=0, Y=2) + P(X=1, Y=2) = 0.31 + 0.11 = 0.42$$
$$P(Y=3) = P(X=0, Y=3) + P(X=1, Y=3) = 0.37 + 0.16 = 0.53$$

Marginal Distributions

Definition

The **marginal distributions** for a jointly distributed pair of r.v.s (X, Y) are the individual distributions for each r.v. If $p_{ij} = P(X = x_i, Y = y_j)$, then we compute the **marginal probabilities** by adding up the joint probabilities over the other index:

$$p_{X,i} = P(X = i) = p_{i1} + p_{12} + \cdots + p_{1m} = \sum_{j=1}^{m} p_{ij}, \qquad 1 \leq i \leq n,$$

and

$$p_{Y,j} = P(Y = j) = p_{1j} + p_{2j} + \cdots + p_{nj} = \sum_{i=1}^{n} p_{ij}, \qquad 1 \leq j \leq m.$$

Independence of Random Variables

We can tell if two r.v.s are independent if their joint probaiblities are the product of their respective marginals.

Definition

Two discrete r.v.s X and Y are **independent** if

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j), \text{ for all } i, j.$$

Rigged Game in New Notation

X and Y are jointly distributed via

Joint Prob.	$Y = y_1 = 1$	$Y = y_2 = 2$	$Y = y_3 = 3$	P_X
$X = x_1 = 0$	$p_{11} = 0.02$	$p_{12} = 0.31$	$p_{13} = 0.37$	$p_{X,1} = 0.7$
$X = x_2 = 1$	$p_{21} = 0.03$	$p_{22} = 0.11$	$p_{23} = 0.16$	$p_{X,2} = 0.3$
P_Y	$p_{Y,1} = 0.05$	$p_{Y,2} = 0.42$	$p_{Y,3} = 0.53$	

CONDITIONAL DISTRIBUTIONS

Definition

Definition

Let X and Y be discrete r.v.s with respective outcomes $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$. The **conditional distribution of X conditional on $Y = y_j$** is given by

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{Y,j}},$$

where p_{ij} is from the joint distribution and $p_{Y,j}$ is from the marginal distribution for Y.

Similarly, the **conditional distribution of Y conditional on $X = x_i$** is given by

$$\mathbb{P}(Y = y_j | X = x_i) = \frac{p_{ij}}{p_{X,i}}$$

Rigged Game Conditional Distributions

Question

Find the conditional distributions of X conditioned on $Y = 1$, $Y = 2$, and $Y = 3$. Then find the conditional distributions of Y conditioned on $X = 0$, then $X = 1$.

If $Y = 1$, we see that

$$P(X = 0|Y = 1) = \frac{P(X = 0, Y = 1)}{P(Y = 1)} = \frac{0.02}{0.05} = 0.4,$$

and

$$P(X = 1|Y = 1) = \frac{P(X = 1, Y = 1)}{P(Y = 1)} = \frac{0.03}{0.05} = 0.6.$$

Conditionals for X conditioned on $Y = 2, 3$

For $Y = 2$ we can compute

$$P(X = 0|Y = 2) = \frac{.31}{.42} \approx 0.7381, \text{ and } P(X = 1|Y = 2) = \frac{.11}{.42} \approx 0.2619,$$

For $Y = 3$ we can compute

$$P(X = 0|Y = 3) = \frac{.37}{.53} \approx 0.6981, \text{ and } P(X = 1|Y = 3) = \frac{.16}{.53} \approx 0.3019,$$

Conditionals for Y conditioned on $X = 0$ and $X = 1$

The conditional distribution of Y conditioned on $X = 0$ can be found via

$$P(Y = 1|X = 0) = \frac{0.02}{0.7} \approx 0.0286,$$
$$P(Y = 2|X = 0) = \frac{0.31}{0.7} \approx 0.4429,$$
$$P(Y = 3|X = 0) = \frac{0.37}{0.7} \approx 0.5286,$$

and conditioned on $X = 1$,

$$P(Y = 1|X = 1) = \frac{0.03}{0.3} \approx 0.1,$$
$$P(Y = 2|X = 1) = \frac{0.11}{0.3} \approx 0.3667,$$
$$P(Y = 3|X = 1) = \frac{0.16}{0.3} \approx 0.5333,$$

Independence implies Conditionals = Marginals

Suppose X and Y are independent and distributed via

$$P(X = 0) = 0.2,\ P(X = 2) = 0.3,\ P(X = 4) = 0.5,$$

and

$$P(Y = 1) = 0.6,\ P(Y = 2) = 0.1,\ P(Y = 3) = 0.3.$$

Since they are independent r.v.s, we have $p_{ij} = p_{X,i}p_{Y,j}$. The joint and marginal probabilties can be displayed as

Joint Probability	$Y = 1$	$Y = 2$	$Y = 3$	P_X
$X = 0$	0.12	0.02	0.06	$p_{X,1} = 0.2$
$X = 2$	0.18	0.03	0.09	$p_{X,2} = 0.3$
$X = 4$	0.3	0.05	0.15	$p_{X,3} = 0.5$
P_Y	$p_{Y,1} = 0.6$	$p_{Y,2} = 0.1$	$p_{Y,3} = 0.3$	

...

Joint Probability	$Y = 1$	$Y = 2$	$Y = 3$	P_X
$X = 0$	0.12	0.02	0.06	$p_{X,1} = 0.2$
$X = 2$	0.18	0.03	0.09	$p_{X,2} = 0.3$
$X = 4$	0.3	0.05	0.15	$p_{X,3} = 0.5$
P_Y	$p_{Y,1} = 0.6$	$p_{Y,2} = 0.1$	$p_{Y,3} = 0.3$	

then since X and Y are independent,

$$P(X = 0|Y = 3) = P(X = 0) = 0.2,\ P(X = 2|Y = 3) = P(X = 2) = 0.3,$$

and

$$P(X = 4|Y = 3) = P(X = 4) = 0.5.$$

Thus the conditional distribution of X conditioned on $Y = 3$ is the same as the marginal distribution of X.

Joint distribution of continuous variables

X, Y continuous random variables with joint CDF

$$F(x, y) = P(X \leq x, Y \leq y).$$

Their **joint PDF** is a function $f(x, y)$:

(i) $f(x, y) \geq 0$ for any $x, y \in \mathbb{R}$.

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

(iii) $P((X, Y) \in A) = \iint\limits_{A} f(x, y) dy dx$ for any region A on $x - y$ plane.

Fundamental theorem of multivariable calculus

- Applying (iii) with $A = \{(X, Y) : X \leq x, Y \leq y\}$,

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v) dv du. \quad (1)$$

- By (1) and the fundamental theorem of calculus,

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

whenever the partial derivatives are defined.

Marginal PDFs of X and Y

- X, Y jointly distributed variables with joint PDF $f(x, y)$.
- The **marginal PDF** f_X of X and f_Y of Y are

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f(x, y) dy \\
f_y(y) &= \int_{-\infty}^{\infty} f(x, y) dx
\end{aligned}
$$

Joint CDF and marginal CDFs

- The **joint CDF** and **marginal PDFs** of X and Y are

$$
\begin{aligned}
F(x,y) &= P(X \le x, Y \le y) \\
F_X(x) &= P(X \le x) \\
F_Y(y) &= P(Y \le y)
\end{aligned}
$$

Concrete formula for joint CDF and marginal PDFs

- If X and Y are discrete with joint PMF $p(x, y)$, then

$$\begin{aligned} F(x, y) &= \sum_{x_i \le x, y_j \le y} p(x_i, y_j), \\ F_X(x) &= \sum_{x_i \le x} p_X(x_i),\ F_Y(y) = \sum_{y_j \le y} p_Y(y_j). \end{aligned}$$

- If X and Y are continuous with joint CDF $f(x, y)$, then

$$\begin{aligned} F(x, y) &= \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v) dv du, \\ F_X(x) &= \int_{-\infty}^{x} f_X(u) du,\ F_Y(y) = \int_{-\infty}^{y} f_Y(v) dv. \end{aligned}$$

Example

Let X and Y be continuous random variables with joint PDF

$$f(x,y) = \begin{cases} \frac{8}{3}x^3 y \text{ if } 0 \le x \le 1,\ 1 \le y \le 2, \\ 0 \text{ otherwise.} \end{cases}$$

Find the marginal PDFs $f_X(x)$ of X and $f_Y(y)$ of Y.

Example solution

Example

Find the joint CDF of X and Y for this example

Example solution

Example

Consider the following function

$$f(x,y) = \begin{cases} a(x^2 + xy) \text{ if } 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 \text{ otherwise} \end{cases} .$$

(a) For what value of a is $f(x,y)$ a joint PDF?
(b) Find marginal PDFs f_X of X and f_Y of Y.
(c) Find the joint CDF $F(x,y)$ of X and Y.
(d) Find $P(X < Y)$.

Example solution

(a) As f is a joint PDF, $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dydx = 1$

$$\begin{aligned}
\int_0^1\int_0^1 a(x^2+xy)dxdy &= 1\\
\int_0^1 \left(a(\frac{x^3}{3}+\frac{x^2}{2}y)\Big|_0^1\right)dy &= 1\\
\int_0^1 a\left(\frac{1}{3}+\frac{1}{2}y\right)dy &= 1\\
a\left(\frac{y}{3}+\frac{y^2}{4}\right)\Big|_0^1 &= 1\\
a &= \frac{12}{7}
\end{aligned}$$

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

- Case 1: $x < 0$ or $x > 1$

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy = \int_{-\infty}^{\infty} 0dy = 0.$$

- Case 2: $0 \le x \le 1$.

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x, y)dy = \int_0^1 \frac{12}{7}(x^2 + xy)dy \\ &= \frac{12}{7}\left(x^2 y + x\frac{y^2}{2}\right)\Big|_0^1 = \frac{12}{7}\left(x^2 + \frac{x}{2}\right). \end{aligned}$$

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

- Conclusion for f_X

$$f_X(x) = \begin{cases} \frac{12}{7}\left(x^2 + \frac{x}{2}\right) & \text{if } 0 \le x \le 1, \\ 0 \text{ otherwise.} \end{cases}$$

- Similarly, we obtain

$$f_Y(y) = \begin{cases} \frac{12}{7}\left(\frac{1}{3} + \frac{y}{2}\right) & \text{if } 0 \le y \le 1, \\ 0 \text{ otherwise.} \end{cases}$$

(c) $F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v)dvdu$

- Case 1: $x < 0$. $f(u, v) = 0$ for any $u \leq x \Rightarrow F(x, y) = 0$.
- Case 2: $0 \leq x \leq 1$. Three subcases concerning y.
  - Subcase 1: $y < 0$. $f(u, v) = 0$ for any $v \leq y \Rightarrow F(x, y) = 0$.
  - Subcase 2: $0 \leq y \leq 1$.

$$\begin{aligned} F(x, y) &= \int_0^x \int_0^y \frac{12}{7}\left(u^2 + uv\right) dvdu = \int_0^x \left(\frac{12}{7}(u^2 v + u\frac{v^2}{2})\Big|_{v=0}^{y}\right) d \\ &= \int_0^x \frac{12}{7}\left(u^2 y + \frac{uy^2}{2}\right) du = \frac{12}{7}\left(\frac{x^3 y}{3} + \frac{x^2 y^2}{4}\right). \end{aligned}$$

  - Subcase 3: $y > 1$.

$$F(x, y) = \int_0^x \int_0^1 \frac{12}{7}\left(u^2 + uv\right) dvdu = \frac{12}{7}\left(\frac{x^3}{3} + \frac{x^2}{4}\right).$$

(c) $F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)dvdu$

- Case 3: $x > 1$. Three subcases concerning y.
  - Subcase 1: $y < 0. f(u,v) = 0$ for any $v \leq y \Rightarrow F(x,y) = 0$.
  - Subcase 2: $0 \leq y \leq 1$.

$$F(x,y) = \int_0^1 \int_0^y \frac{12}{7}\left(u^2 + uv\right) dvdu = \frac{12}{7}\left(\frac{y}{3} + \frac{y^2}{4}\right).$$

  - Subcase 3: $y > 1$.

$$F(x,y) = \int_0^1 \int_0^1 \frac{12}{7}\left(u^2 + uv\right) dvdu = 1,$$

where the last equation follows from property (ii) of $f(x,y)$.

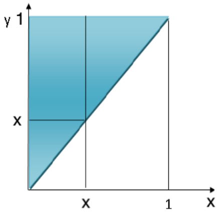
(c) $F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v) dv du$

Conclusion

$$F(x, y) = \begin{cases} 0 \text{ if } x < 0 \text{ or } y < 0, \\ \frac{12}{7}\left(\frac{x^3 y}{3} + \frac{x^2 y^2}{4}\right) \text{ if } 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 1, \\ \frac{12}{7}\left(\frac{x^3}{3} + \frac{x^2}{4}\right) \text{ if } 0 \leq x \leq 1 \text{ and } y > 1, \\ \frac{12}{7}\left(\frac{y}{3} + \frac{y^2}{2}\right) \text{ if } x > 1 \text{ and } 0 \leq y \leq 1, \\ 1 \text{ if } x > 1 \text{ and } y > 1. \end{cases}$$

(d) $P(X < Y)$

- Region $A = \{(x, y) : x < y\}$



- Hence

$$
\begin{aligned}
P(X < Y) &= \iint_A f(x,y)dydx \\
&= \int_0^1 \int_x^1 \frac{12}{7}\left(x^2 + xy\right) dydx
\end{aligned}
$$

(d) $P(X < Y)$

$$
\begin{aligned}
P(X<Y) &= \int_0^1 \int_x^1 \frac{12}{7}\left(x^2+xy\right) dydx \\
&= \int_0^1 \left(\frac{12}{7}(x^2y+\frac{xy^2}{2})\Big|_{y=x}^1 \right) dx \\
&= \int_0^1 \frac{12}{7}\left(x^2+\frac{x}{2}-\frac{3x^3}{2}\right) dx \\
&= \frac{5}{14}.
\end{aligned}
$$

Exercise

Consider the following function.

$$f(x, y) = \begin{cases} ae^{-\lambda y} \text{ if } 0 \le x \le y, \\ 0 \text{ otherwise.} \end{cases}$$

(a) For what value of a is $f(x, y)$ a joint PDF?

(b) Find marginal PDFs of X and Y.

Exercise solution

Exercise solution