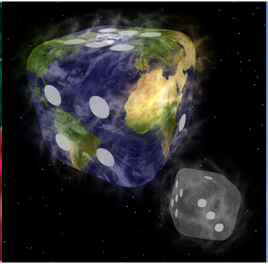
CHAPTER 6: Jointly Distributed Random Variables

Lecturer: Nguyen Minh Quan, PhD
Department of Mathematics
HCMC International University


CONTENTS

1. Joint Distribution Functions
2. Independent Random Variables
3. Sum of independent Random Variables
4. Conditional Distributions: Discrete case and continuous case
5. Joint Probability Distribution of Functions of Random Variables

Introduction

- We want to study two or more random variables at same time.
- We want to describe the relation between them.
- We need to look at their *joint distribution.*

Joint Distribution Function

- Let X and Y be random variables. The joint cumulative distribution function (joint cdf) of X and Y is

$$F_{X,Y}(x,y) = P(X \leq x, Y \leq y).$$

- We can use $F_{X,Y}$ to calculate probability involve X and Y.
- Obtain single cdf of X from joint cdf:

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= P(X \leq x, Y \leq \infty) \\ &= \lim_{y \to \infty} F_{X,Y}(x,y) \end{aligned}$$

- Same for Y: $F_Y(y) = \lim_{x \to \infty} F_{X,Y}(x,y)$

Single cumulative distribution function

Example 1a

Evaluate $P(X > x, Y > y)$.

$$
\begin{aligned}
&P(X > x, Y > y) = \\
&\quad = 1 - P[(X > x, Y > y)^c] \\
&\quad = 1 - P[(X > x)^c \cup (Y > y)^c] \\
&\quad = 1 - P[(X \le x) \cup (Y \le y)] \\
&\quad = 1 - P(X \le x) - P(Y \le y + P(X \le x, Y \le y) \\
&\quad = 1 - F_X(x) - F_Y(y) + F_{X,Y}(x, y).
\end{aligned}
$$

General case:

$$P(x_1 < X \le x_2, y_1 < Y \le y_2) = F(x_2, y_2) + F(x_1, y_1) - F(x_1, y_2) - F(x_2, y_1)$$

Prove it?

Discrete joint probability mass function (pmf)

- X and Y are discrete RVs
- The joint probability mass function (pmf) of X and Y
  $p(x, y) = P(X = x, Y = y)$
- Single pmf

$$p_X(x) = P(X = x) = \sum_y p(x, y)$$

$$p_Y(y) = P(Y = y) = \sum_x p(x, y)$$

Joint probability mass function (pmf)

Example 1b

15% of the families in a community have no children, 20% have 1 child, 35% have 2 children, and 30% have 3. In each family each child is equally likely (independently) to be a boy or a girl. If a family is chosen at random, find the joint pmf of B (the number of boys) and G (the number of girls) in this family.

$$P\{B=0,G=0\} = P\{\text{no children}\} = .15$$

$$\begin{aligned}P\{B=0,G=1\} &= P\{\text{1 girl and total of 1 child}\}\\ &= P\{\text{1 child}\}P\{\text{1 girl}|\text{1 child}\} = (.20)\left(\frac{1}{2}\right)\end{aligned}$$

$$\begin{aligned}P\{B=0,G=2\} &= P\{\text{2 girls and total of 2 children}\}\\ &= P\{\text{2 children}\}P\{\text{2 girls}|\text{2 children}\} = (.35)\left(\frac{1}{2}\right)^2\end{aligned}$$

TABLE 6.2: $P\{B = i, G = j\}$

i \ j	0	1	2	3	Row sum $= P\{B = i\}$
0	.15	.10	.0875	.0375	.3750
1	.10	.175	.1125	0	.3875
2	.0875	.1125	0	0	.2000
3	.0375	0	0	0	.0375
Columnsum $= P\{G = j\}$	.3750	.3875	.2000	.0375	

Joint probability mass function (pmf)

Example 1c

Suppose that 3 balls are randomly selected from an urn containing 3 red, 4 white, and 5 blue balls. If we let X and Y denote, respectively, the number of red and white balls chosen, find the joint probability mass function of X and Y.

$$p(0,0) = \binom{5}{3} \Big/ \binom{12}{3} = \frac{10}{220}$$

$$p(0,1) = \binom{4}{1}\binom{5}{2} \Big/ \binom{12}{3} = \frac{40}{220}$$

$$p(0,2) = \binom{4}{2}\binom{5}{1} \Big/ \binom{12}{3} = \frac{30}{220}$$

$$p(0,3) = \binom{4}{3} \Big/ \binom{12}{3} = \frac{4}{220}$$

$$p(1,0) = \binom{3}{1}\binom{5}{2} \Big/ \binom{12}{3} = \frac{30}{220}$$

$$p(1,1) = \binom{3}{1}\binom{4}{1}\binom{5}{1} \Big/ \binom{12}{3} = \frac{60}{220}$$

$$p(1,2) = \binom{3}{1}\binom{4}{2} \Big/ \binom{12}{3} = \frac{18}{220}$$

$$p(2,0) = \binom{3}{2}\binom{5}{1} \Big/ \binom{12}{3} = \frac{15}{220}$$

$$p(2,1) = \binom{3}{2}\binom{4}{1} \Big/ \binom{12}{3} = \frac{12}{220}$$

$$p(3,0) = \binom{3}{3} \Big/ \binom{12}{3} = \frac{1}{220}$$

Joint probability mass function (pmf)

TABLE 6.1: $P\{X = i, Y = j\}$

$i \backslash j$	0	1	2	3	Row sum $= P\{X = i\}$
0	$\frac{10}{220}$	$\frac{40}{220}$	$\frac{30}{220}$	$\frac{4}{220}$	$\frac{84}{220}$
1	$\frac{30}{220}$	$\frac{60}{220}$	$\frac{18}{220}$	0	$\frac{108}{220}$
2	$\frac{15}{220}$	$\frac{12}{220}$	0	0	$\frac{27}{220}$
3	$\frac{1}{220}$	0	0	0	$\frac{1}{220}$
Column sum $= P\{Y = j\}$	$\frac{56}{220}$	$\frac{112}{220}$	$\frac{48}{220}$	$\frac{4}{220}$	

Joint probability density function (pdf)

Definition

X and Y are *jointly continuous* if there exists function $f(x, y)$ defined on $\mathbb{R}^2$ so that for any Borel subset B

$$P[(X, Y) \in B] = \iint_B f(x, y)dxdy$$

$f(x, y)$ is the joint pdf of X and Y

If A and B are Borel subset of $\mathbb{R}$ then

$$P[(X, Y) \in (A \times B)] = \int_A \int_B f(x, y)dxdy$$

$$F(a, b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f(x, y)dxdy$$

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

Single pdf

- Q: How to get a single PDF from joint PDF?
- If X and Y are jointly continuous then they are continuous with pdf. The single PDFs are:

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

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

Joint PDF and single PDF

Example 1d

The joint pdf of X and Y is

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & 0 < x, y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Compute:
(a) $P(X > 1, Y < 1)$; (b) $P(X < Y)$; (c) $P(X < a)$.

Solution: (a)

$$\begin{aligned} P\{X > 1, Y < 1\} &= \int_0^1 \int_1^\infty 2e^{-x}e^{-2y}\,dx\,dy \\ &= \int_0^1 2e^{-2y}\left(-e^{-x}\Big|_1^\infty\right) dy \\ &= e^{-1}\int_0^1 2e^{-2y}\,dy \\ &= e^{-1}(1 - e^{-2}) \end{aligned}$$

(b)

$$
\begin{aligned}
P(X < Y) &= \iint_{x<y} 2e^{-x}e^{-2y}\,dxdy \\
&= \int_0^{\infty}\int_0^{y} 2e^{-x}e^{-2y}\,dxdy \\
&= 1/3
\end{aligned}
$$

(c)

$$
\begin{aligned}
P(X < a) &= \int_0^{a}\int_0^{\infty} 2e^{-x}e^{-2y}\,dxdy \\
&= 1 - e^{-a}
\end{aligned}
$$

Joint pdf and single PDF

Example 1e

Joint pdf of X and Y

$$f(x,y) = \begin{cases} e^{-(x+y)} & 0 < x, y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Find the pdf of X/Y

Compute the cdf of X/Y:

$$F_{X/Y}(z) = P(X/Y \le z = \iint_{x/y \le z} e^{-(x+y)} dxdy$$

$$= \int_0^\infty \int_0^{zy} e^{-(x+y)} dxdy = 1 - \frac{1}{z+1}$$

Taking derivative of the cdf, we get the pdf of X/Y

$$f(z) = \frac{1}{(z+1)^2}, \quad 0 < z < \infty$$

Independent Random Variables

Independent Random Variables

Definition

Two random variables X and Y are *independent* if for any Borel sets A and B, the events $(X \in A)$ and $(Y \in B)$ are independent.

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

It's very easy to find the joint cdf of independent RVs

$$F(x, y) = F_X(x)F_Y(y)$$

Independent Random Variables

Example 2f

The number of people who enter a post office on a given day is a Poisson RV with parameter λ. Show that if each person is a male with probability p and a female with probability $1 - p$, then the number of males and females entering the post office are independent Poisson RVs with parameters λp and $\lambda(1 - p)$.

Solution

- Let X = number of males, Y = number of females
- $Z = X + Y$: total number of people entering the post office then $Z = Poisson(\lambda)$
- We have

$$P(X = i, Y = j) = \\ P(X = i, Y = j|Z = i + j)P(Z = i + j)$$

- Need to compute $P(X = i, Y = j|Z = i + j)$ and $P(Z = i + j)$.

Example 2f (Cont.)

- Given $(i+j)$ people entering the post office, probability that exactly i males and j females is $Binomial(i+j, p)$

$$P(X=i, Y=j|Z=i+j) =$$
$$= \binom{i+j}{i} p^i (1-p)^j$$

- Z is Poisson with λ, so $P(Z=i+j) = e^{-\lambda}\frac{\lambda^{i+j}}{(i+j)!}$.

$$\begin{aligned} P\{X=i, Y=j\} &= \binom{i+j}{i} p^i (1-p)^j e^{-\lambda} \frac{\lambda^{i+j}}{(i+j)!} \\ &= e^{-\lambda} \frac{(\lambda p)^i}{i!j!} [\lambda(1-p)]^j \\ &= \frac{e^{-\lambda p}(\lambda p)^i}{i!} e^{-\lambda(1-p)} \frac{[\lambda(1-p)]^j}{j!} \end{aligned}$$

Compute the single pmf

- On the other hand:

$$P(X=i)=\sum_j P(X=i,Y=j)=e^{-\lambda p}\frac{(\lambda p)^i}{i!}$$

so X is Poisson(λp). Also,

$$P(Y=j)=\sum_i P(X=i,Y=j)=e^{-\lambda(1-p)}\frac{(\lambda(1-p))^j}{j!}$$

So Y is Poisson$(\lambda(1-p))$.

- Therefore

$$P(X=i,Y=j)=P(X=i)P(Y=j)$$

then X and Y are independent.

Separable joint pdf

- If X and Y are independent then the joint pdf is product of single pdfs.
- Is the reverse true?
- YES!

Proposition

Two continuous (discrete) RVs X and Y are independent if and only if the joint pdf (pmf) can be expressed as

$$f_{X,Y}(x,y) = h(x)g(y)$$

for all x, y.

Separable joint pdf

Example 2g

Given the joint pdf of X and Y is

$$f(x,y) = 24xy$$

for $0 < x < 1, 0 < y < 1, 0 < x + y < 1$ and zero otherwise.
Are X and Y independent?

Solution: Let
$A = \{0 < x < 1, 0 < y < 1,$
$0 < x + y < 1\}$.

Define the indicator function $I_A(x,y) = 1$ if $(x,y) \in A$ and zero otherwise.

Then $f(x,y) = 24xyI_A(x,y)$ not separable. Therefore, X and Y are not independent (or, dependent).

Sum of independent Random Variables

Sum of independent Random Variables

- X and Y continuous and independent
- We can find the cdf of $X + Y$

$$
\begin{aligned}
F_{X+Y}(a) &= P\{X + Y \leq a\} \\
&= \iint\limits_{x+y\leq a} f_X(x) f_Y(y)\,dx\,dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{a-y} f_X(x) f_Y(y)\,dx\,dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{a-y} f_X(x)\,dx f_Y(y)\,dy \\
&= \int_{-\infty}^{\infty} F_X(a - y) f_Y(y)\,dy
\end{aligned}
$$

pdf of sum

$$
\begin{aligned}
f_{X+Y}(a) &= \frac{d}{da}\int_{-\infty}^{\infty} F_X(a - y) f_Y(y)\,dy \\
&= \int_{-\infty}^{\infty} \frac{d}{da} F_X(a - y) f_Y(y)\,dy \\
&= \int_{-\infty}^{\infty} f_X(a - y) f_Y(y)\,dy
\end{aligned}
$$

f_{X+Y} is called the *convolution* of f_X and f_Y.

Sum of normal RVs and Poisson RVs

Sum of normal RVs

If

$$X_i = N(\mu_i, \sigma_i^2), i = 1, \ldots, n$$

are independent, then

$$\sum_i X_i = N(\sum_i \mu_i, \sum_i \sigma_i^2)$$

Sum of Poisson RVs

If X and Y are independent Poisson RVs with λ_1 and λ_2, then $X + Y$ has Poisson distribution with $\lambda_1 + \lambda_2$.

Conditional Distributions: Discrete case and continuous case

Conditional Distributions: Discrete case

The conditional pmf

If X and Y are discrete random variables, we define the conditional probability mass function of X given that $Y = y$, by

$$p_{X|Y}\left(x|\, y\right) = P\left(X = x|\, Y = y\right) = \frac{p\left(x, y\right)}{p_Y\left(y\right)}$$

for all values of y such that $p_Y(y) > 0$.

The conditional pdf

the conditional probability distribution function of X given that $Y = y$ is defined, for all y such that $p_Y(y) > 0$, by

$$F_{X|Y}\left(x|\, y\right) = P\left(X \leqslant x|\, Y = y\right) = \sum_{a \leqslant x} p_{X|Y}\left(a|\, y\right).$$

Conditional Distributions: Discrete case

Example

Suppose that $p(x, y)$, the joint probability mass function of X and Y, is given by $p(0, 0) = 0.4, p(0, 1) = 0.2, p(1, 0) = 0.1, p(1, 1) = 0.3.$
Calculate the conditional probability mass function of X given that Y = 1.

Solution

Note that

$$p_Y(1) = \sum_x p(x, 1) = p(0, 1) + p(1, 1) = 0.5$$

Therefore,

$$p_{X|Y}(0|1) = \frac{p(0,1)}{p_Y(1)} = \frac{2}{5}$$

$$p_{X|Y}(1|1) = \frac{p(1,1)}{p_Y(1)} = \frac{3}{5}$$

Conditional Distributions: Discrete case

Example

If X and Y are independent Poisson random variables with respective parameters λ_1 and λ_2, calculate the conditional distribution of X given that $X + Y = n$.

Hint:

$$P\left(X = k \middle| X + Y = n\right) = \frac{P\left(X = k\right) P\left(Y = n - k\right)}{P\left(X + Y = n\right)}$$

$$P\left(X = k \middle| X + Y = n\right) = \left(\begin{array}{c} n \\ k \end{array} \right) \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^k \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{n-k}$$

Conditional Distributions: Continuous case

Definition

If X and Y have a joint probability density function $f(x, y)$, then the conditional probability density function of X given that $Y = y$ is defined, for all values of y such that $f_Y(y) > 0$, by

$$f_{X|Y}\left(x|\,y\right) = \frac{f\left(x, y\right)}{f_Y\left(y\right)}$$

If X and Y are jointly continuous, then, for any set A

$$P\left(X \in A|\,Y = y\right) = \int\limits_A f_{X|Y}\left(x|\,y\right) dx$$

Conditional Distributions: Continuous case

We define the conditional cumulative distribution function of X given that $Y = y$ by

$$F_{X|Y}(a|\, y) = P(X \leqslant a|\, Y = y) = \int_{-\infty}^{a} f_{X|Y}(x|\, y)\, dx$$

Example

The joint density of X and Y is given by

$$f(x, y) = \begin{cases} \frac{12}{5} x (2 - x - y), 0 < x < 1, 0 < y < 1 \\ 0, \ \text{otherwise} \end{cases}$$

Compute the conditional density of X given that $Y = y$, where $0 < y < 1$.

Solution For $0 < x < 1, 0 < y < 1$, we have

$$f_{X|Y}(x|\, y) = \frac{f(x, y)}{f_Y(y)} = \frac{f(x, y)}{\int_0^1 \frac{12}{5} x (2 - x - y)\, dx} = \frac{6x(2 - x - y)}{4 - 3y}$$

Conditional Distributions: Continuous case

Example

The joint density of X and Y is given by

$$f(x,y) = \begin{cases} \frac{e^{-x/y}e^{-y}}{y}, 0 < x < \infty, 0 < y < \infty \\ 0,\ \text{otherwise} \end{cases}$$

Find $P(X > 1 | Y = y)$.

Solution

We first obtain the conditional density of X given that $Y = y$

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{e^{-x/y}e^{-y}/y}{e^{-y}\int_0^\infty (1/y)e^{-x/y}dx} = \frac{e^{-x/y}}{y}$$

$$P(X > 1 | Y = y) = \int_1^\infty (1/y)e^{-x/y}dx = e^{-1/y}$$

Joint Probability Distribution of Functions of Random Variables

- Let X_1 and X_2 be jointly continuous random variables with joint probability density function $f_{X1,X2}$.
- Suppose that $Y_1 = g_1(X_1, X_2)$ and $Y_2 = g_2(X_1, X_2)$ for some functions g_1 and g_2. We want to obtain the joint distribution of the random variables Y_1 and Y_2.
- Assume:
  1. The equations $y_1 = g_1(x_1, x_2)$ and $y_2 = g_2(x_1, x_2)$ can be uniquely solved for x_1 and x_2: $x_1 = h_1(y_1, y_2)$, $x_2 = h_2(y_1, y_2)$.
  2. The functions g_1 and g_2 have continuous partial derivatives at all points (x_1, x_2) and are such that

$$J(x_1, x_2) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{vmatrix} \neq 0$$

Joint Probability Distribution of Functions of Random Variables

- Let X_1 and X_2 be jointly continuous random variables with joint probability density function $f_{X1,X2}$.
- Suppose that $Y_1 = g_1(X_1, X_2)$ and $Y_2 = g_2(X_1, X_2)$ for some functions g_1 and g_2. We want to obtain the joint distribution of the random variables Y_1 and Y_2.
- Conditions:
  1. The equations $y_1 = g_1(x_1, x_2)$ and $y_2 = g_2(x_1, x_2)$ can be uniquely solved for x_1 and x_2: $x_1 = h_1(y_1, y_2)$, $x_2 = h_2(y_1, y_2)$.
  2. The functions g_1 and g_2 have continuous partial derivatives at all points (x_1, x_2) and are such that

$$J(x_1, x_2) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{vmatrix} \neq 0$$

Joint Probability Distribution of Functions of Random Variables

Theorem

Under these two conditions, it can be shown that the random variables Y_1 and Y_2 are jointly continuous with joint density function given by

$$f_{Y_1 Y_2}(y_1, y_2) = f_{X_1, X_2}(x_1, x_2)|J(x_1, x_2)|^{-1},$$

where $x_1 = h_1(y_1, y_2)$, $x_2 = h_2(y_1, y_2)$.

Joint Probability Distribution of Functions of Random Variables

Example

Let X_1 and X_2 be jointly continuous random variables with probability density function f_{X_1,X_2}. Let $Y_1 = X_1 + X_2$, $Y_2 = X_1 - X_2$. Find the joint density function of Y_1 and Y_2 in terms of f_{X_1,X_2}.

We have

$$J(x_1, x_2) = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2$$

and $x_1 = (y_1 + y_2)/2, x_2 = (y_1 - y_2)/2$. It follows

$$f_{Y_1,Y_2}(y_1, y_2) = \frac{1}{2} f_{X_1,X_2}\left(\frac{y_1 + y_2}{2}, \frac{y_1 - y_2}{2}\right)$$

-END OF CHAPTER 6-