# Chapter 4: Multiple Random Variables
# Lecture 14: Joint and marginal distributions

So far we focus on the distribution of a single random variable.

In applications we need to consider a set of random variables jointly.

In some cases we study relationships among random variables.

## Definition 4.1.1.

For an integer $n$, $X = (X_1, ..., X_n)$ is called an $n$-dimensional random vector iff each $X_i$ is a random variable.

## Joint cdf's

The joint cdf of an $n$-dimensional random vector $X$ is a function $F_X$ in $\mathscr{R}^n$ such that

$$
\begin{aligned}
F_X(x_1, ..., x_n) &= P\left(\bigcap_{i=1}^{n} \{X_i \le x_i\}\right) & x_i \in \mathscr{R} \\
&= P(X_1 \le x_1, ..., X_n \le x_n) & i = 1, ..., n
\end{aligned}
$$

## Properties of cdf's

- $F_X$ is nondecreasing and right-continuous in any of its $n$ arguments.
- For any $i = 1, ..., n$ and fixed $x_1, ..., x_{i-1}, x_{i+1}, ..., x_n$,

$$\lim_{x_i \to -\infty} F(x_1, ..., x_{i-1}, x_i, x_{i+1}, ..., x_n) = 0$$

$$\lim_{x_i \to \infty,\ i=1,...,n} F(x_1, ..., x_n) = 1$$

## Marginal cdf's

For a random vector $X = (X_1, ..., X_n)$ and any $i$, the cdf of $X_i$ is called the marginal cdf of $X_i$ and is equal to

$$F_{X_i}(x_i) = \lim_{x_j \to \infty,\ j=1,...,i-1,i+1,...,n} F_X(x_1, ..., x_n)$$

Knowing the joint cdf $F_X$ we can obtain $n$ marginal cdf's, but in general, knowing $F_{X_1}, ..., F_{X_n}$ is not enough to determine the joint cdf $F_X$.

Similar to the univariate case, we mainly consider two types of random vectors, discrete random vectors and continuous random vectors.

## Properties of cdf's

- $F_X$ is nondecreasing and right-continuous in any of its $n$ arguments.
- For any $i = 1, ..., n$ and fixed $x_1, ..., x_{i-1}, x_{i+1}, ..., x_n$,

$$\lim_{x_i \to -\infty} F(x_1, ..., x_{i-1}, x_i, x_{i+1}, ..., x_n) = 0$$

$$\lim_{x_i \to \infty,\ i=1,...,n} F(x_1, ..., x_n) = 1$$

## Marginal cdf's

For a random vector $X = (X_1, ..., X_n)$ and any $i$, the cdf of $X_i$ is called the marginal cdf of $X_i$ and is equal to

$$F_{X_i}(x_i) = \lim_{x_j \to \infty,\ j=1,...,i-1,i+1,...,n} F_X(x_1, ..., x_n)$$

Knowing the joint cdf $F_X$ we can obtain $n$ marginal cdf's, but in general, knowing $F_{X_1}, ..., F_{X_n}$ is not enough to determine the joint cdf $F_X$.

Similar to the univariate case, we mainly consider two types of random vectors, discrete random vectors and continuous random vectors.

## Properties of cdf's

- $F_X$ is nondecreasing and right-continuous in any of its $n$ arguments.
- For any $i = 1, ..., n$ and fixed $x_1, ..., x_{i-1}, x_{i+1}, ..., x_n$,

$$\lim_{x_i \to -\infty} F(x_1, ..., x_{i-1}, x_i, x_{i+1}, ..., x_n) = 0$$

$$\lim_{x_i \to \infty,\ i=1,...,n} F(x_1, ..., x_n) = 1$$

## Marginal cdf's

For a random vector $X = (X_1, ..., X_n)$ and any $i$, the cdf of $X_i$ is called the marginal cdf of $X_i$ and is equal to

$$F_{X_i}(x_i) = \lim_{x_j \to \infty,\ j=1,...,i-1,i+1,...,n} F_X(x_1, ..., x_n)$$

Knowing the joint cdf $F_X$ we can obtain $n$ marginal cdf's, but in general, knowing $F_{X_1}, ..., F_{X_n}$ is not enough to determine the joint cdf $F_X$.

Similar to the univariate case, we mainly consider two types of random vectors, discrete random vectors and continuous random vectors.

## Definition 4.1.3 (Discrete joint pmf)

A random vector $X = (X_1,...,X_n)$ is discrete iff each $X_i$ is discrete. The joint pmf of $X$ is

$$f_X(x_1,...,x_n) = P(X_1 = x_1,...,X_n = x_n),$$

which is positive only for countably many $(x_1,...,x_n) \in \mathscr{R}^n$.

- For any event $A \subset \mathscr{R}^n$,

$$P(X \in A) = \sum_{(x_1,...,x_n)\in A} f_X(x_1,...,x_n)$$

- For any $i$, the marginal pmf of $X_i$ is

$$f_{X_i}(x_i) = \sum_{x_1,...,x_{i-1},x_{i+1},...,x_n} f_X(x_1,...,x_n)$$

- For any function $g(x_1,...,x_n)$, the expected value of $g(X_1,...,X_n)$ is

$$E[g(X_1,...,X_n)] = \sum_{x_1,...,x_n} g(x_1,...,x_n) f_X(x_1,...,x_n)$$

- If $n = 2$ and each $X_i$ takes finitely many values, then the joint and marginal pmf's can be listed in a $2 \times 2$ table.

## Definition 4.1.3 (Discrete joint pmf)

A random vector $X = (X_1, ..., X_n)$ is discrete iff each $X_i$ is discrete.
The joint pmf of $X$ is

$$f_X(x_1, ..., x_n) = P(X_1 = x_1, ..., X_n = x_n),$$

which is positive only for countably many $(x_1, ..., x_n) \in \mathscr{R}^n$.

- For any event $A \subset \mathscr{R}^n$,

$$P(X \in A) = \sum_{(x_1,...,x_n) \in A} f_X(x_1, ..., x_n)$$

- For any $i$, the marginal pmf of $X_i$ is

$$f_{X_i}(x_i) = \sum_{x_1,...,x_{i-1},x_{i+1},...,x_n} f_X(x_1, ..., x_n)$$

- For any function $g(x_1, ..., x_n)$, the expected value of $g(X_1, ..., X_n)$ is

$$E[g(X_1, ..., X_n)] = \sum_{x_1,...,x_n} g(x_1, ..., x_n) f_X(x_1, ..., x_n)$$

- If $n = 2$ and each $X_i$ takes finitely many values, then the joint and marginal pmf's can be listed in a $2 \times 2$ table.

## Values and joint and marginal pmf's of a 2-dimensional random vector

| value | $x_{21}$ | $x_{22}$ | $\cdots$ | $x_{2c}$ | marginal |
|---|---|---|---|---|---|
| $x_{11}$ | $p_{11}$ | $p_{12}$ | $\cdots$ | $p_{1c}$ | $p_{1\cdot}$ |
| $x_{12}$ | $p_{21}$ | $p_{22}$ | $\cdots$ | $p_{2c}$ | $p_{2\cdot}$ |
| $\cdots$ | $\cdots$ | $\cdots$ | $\cdots$ | $\cdots$ | $\cdots$ |
| $x_{1r}$ | $p_{r1}$ | $p_{r2}$ | $\cdots$ | $p_{rc}$ | $p_{r\cdot}$ |
| marginal | $p_{\cdot 1}$ | $p_{\cdot 2}$ | $\cdots$ | $p_{\cdot c}$ | 1 |

$$p_{ij} = P(X_1 = i, X_2 = j),\ p_{i\cdot} = P(X_1 = i),\ p_{\cdot j}(X_2 = j).$$

## Example: Multinomial distribution

An experiment has $r$ possible outcomes $A_1, ..., A_r$ with $P(A_i) = p_i$, $i = 1, ..., r$, $p_1 + \cdots + p_r = 1$.

We independently repeat the experiment $n$ times.

If $X_i$ is the number of times $A_i$ is the result in $n$ experiments, $i = 1, ..., r$, then $X = (X_1, ..., X_r)$ has joint pmf

$$P(X_1 = x_1, ..., X_r = x_r) = \frac{n!}{x_1! \cdots x_r!} p_1^{x_1} \cdots p_r^{x_r} \qquad 0 \le x_j,\ \sum_{i=1}^{r} x_i = n$$

Values and joint and marginal pmf's of a 2-dimensional random vector

| value | $x_{21}$ | $x_{22}$ | $\cdots$ | $x_{2c}$ | marginal |
|---|---|---|---|---|---|
| $x_{11}$ | $p_{11}$ | $p_{12}$ | $\cdots$ | $p_{1c}$ | $p_{1\cdot}$ |
| $x_{12}$ | $p_{21}$ | $p_{22}$ | $\cdots$ | $p_{2c}$ | $p_{2\cdot}$ |
| $\cdots$ | $\cdots$ | $\cdots$ | $\cdots$ | $\cdots$ | $\cdots$ |
| $x_{1r}$ | $p_{r1}$ | $p_{r2}$ | $\cdots$ | $p_{rc}$ | $p_{r\cdot}$ |
| marginal | $p_{\cdot 1}$ | $p_{\cdot 2}$ | $\cdots$ | $p_{\cdot c}$ | 1 |

$$p_{ij} = P(X_1 = i, X_2 = j),\ p_{i\cdot} = P(X_1 = i),\ p_{\cdot j}(X_2 = j).$$

## Example: Multinomial distribution

An experiment has $r$ possible outcomes $A_1, ..., A_r$ with $P(A_i) = p_i$, $i = 1, ..., r$, $p_1 + \cdots + p_r = 1$.
We independently repeat the experiment $n$ times.
If $X_i$ is the number of times $A_i$ is the result in $n$ experiments, $i = 1, ..., r$, then $X = (X_1, ..., X_r)$ has joint pmf

$$P(X_1 = x_1, ..., X_r = x_r) = \frac{n!}{x_1! \cdots x_r!} p_1^{x_1} \cdots p_r^{x_r} \qquad 0 \le x_j,\ \sum_{i=1}^{r} x_i = n$$

## Multivariate hypergeometric distribution

A bag contains $N$ balls with $r$ different colors.
$N_i =$ the number of balls for color $i$, $N_1 + \cdots + N_r = N$.
We randomly select $n$ balls from the bag.
If $X_i$ is the number of selected balls having color $i$, $i = 1, ..., r$, then $X = (X_1, ..., X_r)$ has joint pmf

$$P(X_1 = x_1, ..., X_r = x_r) = \frac{\binom{N_1}{x_1} \cdots \binom{N_r}{x_r}}{\binom{N}{n}} \qquad 0 \le x_j, \ \sum_{i=1}^{r} x_i = n$$

What are the marginal pmf's for the multinomial and multivariate hypergeometric distributions?

## Definition 4.1.10 (continuous pdf)

A random vector $X = (X_1, ..., X_n)$ has a continuous joint pdf if there exists a nonnegative function $f_X$ on $\mathscr{R}^n$ such that

$$P(X \in A) = \int \cdots \int_{(x_1,...,x_n) \in A} f_X(x_1, ..., x_n) dx_1 \cdots dx_n$$

## Multivariate hypergeometric distribution

A bag contains $N$ balls with $r$ different colors.
$N_i =$ the number of balls for color $i$, $N_1 + \cdots + N_r = N$.
We randomly select $n$ balls from the bag.
If $X_i$ is the number of selected balls having color $i$, $i = 1, ..., r$, then $X = (X_1, ..., X_r)$ has joint pmf

$$P(X_1 = x_1, ..., X_r = x_r) = \frac{\binom{N_1}{x_1} \cdots \binom{N_r}{x_r}}{\binom{N}{n}} \qquad 0 \leq x_j, \ \sum_{i=1}^{r} x_i = n$$

What are the marginal pmf's for the multinomial and multivariate hypergeometric distributions?

## Definition 4.1.10 (continuous pdf)

A random vector $X = (X_1, ..., X_n)$ has a continuous joint pdf if there exists a nonnegative function $f_X$ on $\mathscr{R}^n$ such that

$$P(X \in A) = \int \cdots \int_{(x_1, ..., x_n) \in A} f_X(x_1, ..., x_n) dx_1 \cdots dx_n$$

## Multivariate hypergeometric distribution

A bag contains $N$ balls with $r$ different colors.
$N_i =$ the number of balls for color $i$, $N_1 + \cdots + N_r = N$.
We randomly select $n$ balls from the bag.
If $X_i$ is the number of selected balls having color $i$, $i = 1, ..., r$, then $X = (X_1, ..., X_r)$ has joint pmf

$$P(X_1 = x_1, ..., X_r = x_r) = \frac{\binom{N_1}{x_1} \cdots \binom{N_r}{x_r}}{\binom{N}{n}} \qquad 0 \le x_j, \ \sum_{i=1}^{r} x_i = n$$

What are the marginal pmf's for the multinomial and multivariate hypergeometric distributions?

## Definition 4.1.10 (continuous pdf)

A random vector $X = (X_1, ..., X_n)$ has a continuous joint pdf if there exists a nonnegative function $f_X$ on $\mathscr{R}^n$ such that

$$P(X \in A) = \int \cdots \int_{(x_1, ..., x_n) \in A} f_X(x_1, ..., x_n) dx_1 \cdots dx_n$$

- If $X=(X_1,...,X_n)$ has joint pdf $f_X$, a short notation is

$$P(X \in A) = \int_A f_X(x)dx \qquad x=(x_1,...,x_n)$$

- If the joint pdf $f_X$ exists, then

$$F_X(x_1,...,x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_X(t_1,...,t_n)dt_1 \cdots dt_n$$

- If the joint cdf $F_X$ is differentiable, then the joint $f_X$ exists and

$$f_X(x_1,...,x_n) = \frac{\partial^n F_X(x_1,...,x_n)}{\partial x_1 \cdots \partial x_n}, \qquad (x_1,...,x_n) \in \mathscr{R}^n$$

- If the joint pdf $f_X$ exists, then the $i$th margianl pdf $f_{X_i}$ exists and

$$f_{X_i}(x_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(x_1,...,x_n)dx_1 \cdots dx_{i-1}dx_{i+1} \cdots dx_n$$

- If the joint pdf $f_X$ exists, then for any function $g(x_1,...,x_n)$, the expected value of $g(X_1,...,X_n)$ is

$$E[g(X_1,...,X_n)] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1,...,x_n)f_X(x_1,...,x_n)dx_1 \cdots dx_n$$

## Example

Suppose that a 2-dimensional random vector $(X, Y)$ has pdf

$$f(x,y) = \begin{cases} Ce^{-(2x+3y)} & x \geq 0, y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- What should $C$ be? Since

$$\begin{aligned} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy &= \int_0^{\infty}\int_0^{\infty} Ce^{-(2x+3y)}dxdy \\ &= C\int_0^{\infty} e^{-2x}dx \int_0^{\infty} e^{-3y}dy \\ &= C\left(\frac{e^{-2x}}{2}\Big|_{\infty}^{0} \frac{e^{-3y}}{3}\Big|_{\infty}^{0}\right) = \frac{C}{6} \end{aligned}$$

  must be 1, we obtain that $C = 6$.

- The joint cdf of $(X, Y)$ is

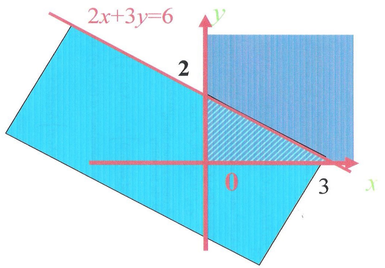
$$F(x,y) = \begin{cases} (1-e^{-2x})(1-e^{-3y}) & x \geq 0, y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

This is because, for $x > 0$ and $y > 0$,

$$\begin{aligned} F(x,y) &= \int_0^\infty \int_0^\infty 6e^{-(2x+3y)} dxdy = 6 \int_0^\infty e^{-2x} dx \int_0^\infty 6e^{-3y} dy \\ &= (1 - e^{-2x})(1 - e^{-3y}) \end{aligned}$$

- Calculate $P(2X + 3Y \leq 6)$.
  It is not convenient to use the joint cdf.
  Using the joint pdf, we obtain

$$P(2X + 3Y \leq 6) = \int_{2x+3y \leq 6} f(x,y) dxdy$$



$$
\begin{aligned}
&= 6\int_0^3 \left[\int_0^{(6-2x)/3} e^{-(2x+3y)}dy\right]dx = 6\int_0^3 e^{-2x}\left[-\frac{e^{-3y}}{3}\bigg|_0^{(6-2x)/3}\right]dx \\
&= 6\int_0^3 e^{-2x}\left[\frac{1}{3}-\frac{e^{-(6-2x)}}{3}\right]dx = 2\int_0^3 (e^{-2x}-e^{-6})dx \\
&= -e^{-2x}\bigg|_0^3 - 2e^{-6}\times 3 = 1-e^{-6}-6e^{-6} \\
&= 1-7e^{-6}
\end{aligned}
$$

## Example 4.1.12.

Suppose that a 2-dimensional random vector $(X,Y)$ has pdf

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

We want to calculate $P(X+Y \geq 1)$.

$$
\begin{aligned}
&= 6\int_0^3 \left[\int_0^{(6-2x)/3} e^{-(2x+3y)}dy\right]dx = 6\int_0^3 e^{-2x}\left[-\frac{e^{-3y}}{3}\Big|_0^{(6-2x)/3}\right]dx \\
&= 6\int_0^3 e^{-2x}\left[\frac{1}{3}-\frac{e^{-(6-2x)}}{3}\right]dx = 2\int_0^3 (e^{-2x}-e^{-6})dx \\
&= -e^{-2x}\Big|_0^3 - 2e^{-6}\times 3 = 1-e^{-6}-6e^{-6} \\
&= 1-7e^{-6}
\end{aligned}
$$

## Example 4.1.12.

Suppose that a 2-dimensional random vector $(X, Y)$ has pdf

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

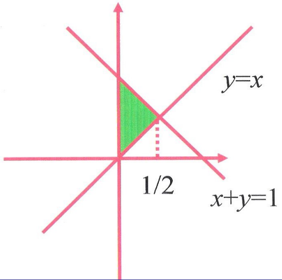
We want to calculate $P(X+Y \geq 1)$.

Note that

$$\begin{aligned}
P(X+Y<1) &= \int_{x+y<1} f(x,y)dxdy = \int_0^{1/2}\left(\int_x^{1-x} e^{-y}dy\right)dx \\
&= \int_0^{1/2}(e^{-x}-e^{-(1-x)})dx = 1-e^{-1/2}-e^{-1/2}+e^{-1} \\
&= 1+e^{-1}-2e^{-1/2}
\end{aligned}$$

Hence,

$$P(X+Y\geq 1) = 1-P(X+Y<1) = 2e^{-1/2}-e^{-1}$$



## Example

Let

$$f(x,y) = \begin{cases} \frac{x^{\alpha-1}(y-x)^{\beta-1}e^{-y}}{\Gamma(\alpha)\Gamma(\beta)} & 0 < x < y \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha > 0$ and $\beta > 0$ are constants.

We want to show this is a pdf and find its two marginal pdf's.

For $x > 0$, consider

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x,y)dy = \int_x^{\infty} \frac{x^{\alpha-1}(y-x)^{\beta-1}e^{-y}}{\Gamma(\alpha)\Gamma(\beta)} dy \\ &= \frac{x^{\alpha-1}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^{\infty} u^{\beta-1} e^{-(u+x)} du \qquad y-x=u \\ &= \frac{x^{\alpha-1}}{\Gamma(\alpha)} e^{-x} \end{aligned}$$

This is the pdf of $Gamma(\alpha,1)$, which also shows that $f(x,y)$ is a pdf.

The other marginal pdf is, for $y > 0$,

$$
\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)dx = \int_0^y \frac{x^{\alpha-1}(y-x)^{\beta-1}e^{-y}}{\Gamma(\alpha)\Gamma(\beta)}dx \\
&= \frac{e^{-y}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 (ty)^{\alpha-1}(y-ty)^{\beta-1} y dt \qquad\qquad x = ty \\
&= \frac{y^{\alpha+\beta-1}e^{-y}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt = \frac{y^{\alpha+\beta-1}e^{-y}}{\Gamma(\alpha+\beta)}
\end{aligned}
$$

which is the pdf of $Gamma(\alpha+\beta,1)$.

## Example

Let $F(x)$ and $G(y)$ be univariate cdf's with pdf's $f(x)$ and $g(y)$, respectively, and $\alpha$ be a constant with $|\alpha| \le 1$.
We want to show that

$$H(x,y) = F(x)G(y)\{1+\alpha[1-F(x)][1-G(y)]\} \qquad (x,y) \in \mathscr{R}^2$$

is a 2-dimensional cdf with pdf

$$h(x,y) = f(x)g(y)\{1+\alpha[2F(x)-1][2G(y)-1]\} \qquad (x,y) \in \mathscr{R}^2$$

Since

$$|\alpha[2F(x)-1][2G(y)-1]| \le |[2F(x)-1][2G(y)-1]| \le 1$$

we know that $h(x,y) \ge 0$.
For any $(x,y) \in \mathscr{R}^2$,

$$\begin{aligned}
&\int_{-\infty}^{x}\int_{-\infty}^{y} h(u,v)dvdu \\
= &\int_{-\infty}^{x}\int_{-\infty}^{y} f(u)g(v)\{1+\alpha[2F(u)-1][2G(v)-1]\}dvdu \\
= &\ F(x)G(y)+\alpha\int_{-\infty}^{x} f(u)[2F(u)-1]du\int_{-\infty}^{y} g(v)[2G(v)-1]dv \\
= &\ F(x)G(y)+\alpha[2F^2(x)-F(x)][G^2(y)-G(y)] \\
= &\ F(x)G(y)\{1+\alpha[1-F(x)][1-G(y)]\} \\
= &\ H(x,y)
\end{aligned}$$

This shows that $h(x,y)$ is a pdf of $H(x,y)$, if $H(x,y)$ is a cdf.
$H(x,y)$ is indeed a cdf, because

$$\lim_{x\to\infty,y\to\infty} H(x,y) = \lim_{x\to\infty,y\to\infty} F(x)G(y) = 1$$