## JOINT DISTRIBUTIONS

## Outlines

-Discrete/Continuous Random Bivariate Variables

- Joint Probability Distributions
- Marginal Probability Distributions
- Conditional Probability Distributions
- Independence, Covariance and Correlation
-Random vector


## Introduction

## Bivariate Frequency Distributions

- For example, suppose you throw two coins, $X$ and $Y$, simultaneously and record the outcome as an ordered pair of values. Imagine that you threw the coin 8 times, and observed the following ( $1=$ Head, $0=$ Tail)

| $(\boldsymbol{X}, \boldsymbol{Y})$ | $\boldsymbol{f}$ |
| :---: | :---: |
| $(1,1)$ | 2 |
| $(1,0)$ | 2 |
| $(0,1)$ | 2 |
| $(0,0)$ | 2 |

- To graph the bivariate distribution, you need a 3 dimensional plot, although this can be drawn in perspective in 2 dimensions


## Joint distribution of bivariate random variables

$\square$ In general, if X and Y are two random variables, the probability distribution that defines their simultaneous behavior is called a joint probability distribution.
$\square$ For example: X : the length of one dimension of an injection-molded part, and Y : the length of another dimension. We might be interested in

$$
\mathrm{P}(2.95 \leq \mathrm{X} \leq 3.05 \text { and } 7.60 \leq \mathrm{Y} \leq 7.80)
$$

## Discrete case

Definition: If $X$ and $Y$ are discrete RV's, then ( $X, Y$ )
is called a jointly discrete bivariate RV.

The joint (or bivariate) pmf is

$$
f(x, y)=\operatorname{Pr}(X=x, Y=y)
$$

Properties:
(1) $0 \leq f(x, y) \leq 1$.
(2) $\Sigma_{x} \Sigma_{y} f(x, y)=1$.
(3) $A \subseteq \Re^{2} \Rightarrow \operatorname{Pr}((X, Y) \in A)=\sum \sum_{(x, y) \in A} f(x, y)$.

## Continuous case

Definition: If $X$ and $Y$ are cts RV's, then $(X, Y)$ is a jointly cts $\mathbf{R V}$ if there exists a function $f(x, y)$ such that

I
(1) $f(x, y) \geq 0, \forall x, y$.
(2) $\iint_{\Re^{2}} f(x, y) d x d y=1$.
(3) $\operatorname{Pr}(A)=\operatorname{Pr}((X, Y) \in A)=\iint_{A} f(x, y) d x d y$. between $A$ and $\mathbf{f}(\mathbf{x}, \mathbf{y})$

## Marginal Probability Distribution

## Discrete case

Definition: If $X$ and $Y$ are jointly discrete, then the marginal pmf's of $X$ and $Y$ are, respectively,

$$
f_{X}(x)=\sum_{y} f(x, y)
$$

and

## Continuous case

$$
f_{Y}(y)=\sum_{x} f(x, y)
$$

Definition: If $X$ and $Y$ are jointly continuous random variable, then the marginal pdf's of X and Y are, respectively;

$$
f_{X}(x)=\int_{y} f(x, y) d y
$$

and

$$
f_{Y}(y)=\int_{x} f(x, y) d x
$$

## General Example for discrete bivariate random variables

If X and Y are two discrete random variables ,their joint distribution may be represented by a formula or a table below:

|  | Y1 | Y2 | -• | Yn | Marginal pdf of $\mathbf{X}$ g(X) |
| :---: | :---: | :---: | :---: | :---: | :---: |
| X1 | $\mathrm{f}(\mathrm{x} 1, \mathrm{y} 1)$ | $\mathrm{f}(\mathrm{x} 1, \mathrm{y} 2)$ | .... | $\mathbf{f}(\mathbf{x 1 , y n})$ | g(x1) |
| X2 | $\mathbf{f}(\mathbf{x} 2, \mathrm{y} 1)$ | $\mathrm{f}(\mathrm{x} 2, \mathrm{y} 2$ | ... | $\mathbf{f}(\mathbf{x} 2, \mathrm{yn})$ | $\mathrm{g}(\mathrm{x} 2)$ |
| ..... | .... | .... | .... | .... | ... |
| Xm | $\mathbf{f}(\mathrm{xm}, \mathrm{y} 1)$ | $\mathrm{f}(\mathrm{xm}, \mathrm{y} 2$ | .... | $\mathbf{f}(\mathbf{x} 2, \mathrm{yn})$ | $\mathrm{g}(\mathrm{xm})$ |
| $\begin{aligned} & \text { Marginal } \\ & \text { pdf of } Y \\ & h(y) \end{aligned}$ | h(y1) | h(y2) | ... | h(yn) | Total $=1$ |

## Discrete case example

|  | $\mathrm{X}=1$ | $\mathrm{X}=3$ | $\mathrm{X}=7$ | $\mathrm{P}(\mathrm{Y}=\mathrm{y})$ |
| :---: | :---: | :---: | :---: | :---: |
| $\mathrm{Y}=10$ | $2 / 7$ | $1 / 7$ | 0 | $9 / 21$ |
| $\mathrm{Y}=15$ | $1 / 3$ | $1 / 21$ | 0 | $8 / 21$ |
| $\mathrm{Y}=20$ | $1 / 7$ | 0 | $1 / 21$ | $4 / 21$ |
| $\mathrm{P}(\mathrm{X}=\mathrm{x})$ | $16 / 21$ | $4 / 21$ | $1 / 21$ | 1 |

The marginal pmf's of $X$ and $Y$ respectively, are given by:
$P(X=x)=f_{X}(x)= \begin{cases}\frac{16}{21}, & x=1 \\ \frac{4}{21}, & x=3 \\ \frac{1}{21}, & x=7\end{cases}$
where $16 / 21+4 / 21+1 / 21=1$
$\mathbf{P}(\mathbf{X}=2)=$ ?

$$
P(Y=15)=?
$$

$$
P(Y=y)=f_{Y}(y)= \begin{cases}\frac{9}{21}, & y=10 \\ \frac{8}{21}, & y=15 \\ \frac{4}{21}, & y=20\end{cases}
$$

Also,
$9 / 21+8 / 21+4 / 21=1$

## Discrete case example (Read)

Example: 3 sox in a box (numbered 1,2,3). Draw 2 sox at random w/o replacement. $X=\#$ of first sock, $Y=\#$ of second sock. The joint pmf $f(x, y)$ is

|  | $X=1$ | $X=2$ | $X=3$ | $\operatorname{Pr}(Y=y)$ |
| :---: | :---: | :---: | :---: | :---: |
| $Y=1$ | 0 | $1 / 6$ | $1 / 6$ | $1 / 3$ |
| $Y=2$ | $1 / 6$ | 0 | $1 / 6$ | $1 / 3$ |
| $Y=3$ | $1 / 6$ | $1 / 6$ | 0 | $1 / 3$ |
| $\operatorname{Pr}(X=x)$ | $1 / 3$ | $1 / 3$ | $1 / 3$ | 1 |

$\operatorname{Pr}(X=x)$ is the "marginal" distribution of $X$.
$\operatorname{Pr}(Y=y)$ is the "marginal" distribution of $Y$.

By the law of total probability,

$$
\operatorname{Pr}(X=1)=\sum_{y=1}^{3} \operatorname{Pr}(X=1, Y=y)=1 / 3
$$

In addition,

$$
\begin{aligned}
& \operatorname{Pr}(X \geq 2, Y \geq 2) \\
& \quad=\sum_{x \geq 2} \sum_{y \geq 2} f(x, y) \\
& =f(2,2)+f(2,3)+f(3,2)+f(3,3) \\
& =0+1 / 6+1 / 6+0=1 / 3 .
\end{aligned}
$$

## Continuous case example: Read

Example: Suppose that

$$
f(x, y)=\left\{\begin{array}{lc}
4 x y & \text { if } 0 \leq x \leq 1,0 \leq y \leq 1 \\
0 & \text { otherwise }
\end{array}\right.
$$

Find the prob (volume) of the region $0 \leq y \leq 1-x^{2}$.

$$
\begin{aligned}
V & =\int_{0}^{1} \int_{0}^{1-x^{2}} 4 x y d y d x \\
& =\int_{0}^{1} \int_{0}^{\sqrt{1-y}} 4 x y d x d y \\
& =1 / 3
\end{aligned}
$$

## Example of joint density for continuous r.v.'s

- Let the joint density of X and Y be

$$
\mathrm{f}(\mathrm{x}, \mathrm{y})=\left\{\begin{array}{cl}
2 \mathrm{e}^{-\mathrm{x}} \mathrm{e}^{-2 \mathrm{y}}, & 0<\mathrm{x}<\infty, \quad 0<\mathrm{y}<\infty \\
0, & \text { otherwise }
\end{array}\right.
$$

- Prove that
(1) $\mathrm{P}\{\mathrm{X}>1, \mathrm{Y}<1\}=\mathrm{e}^{-1}\left(1-\mathrm{e}^{-2}\right)$
(2) $\mathrm{P}\{\mathrm{X}<\mathrm{Y}\}=1 / 3$
(3) $\mathrm{F}_{\mathrm{X}}(\mathrm{a})=1-\mathrm{e}^{-\mathrm{a}}, \mathrm{a}>0$, and 0 otherwise.

Note: Going from cdf's to pdf's (continuous case).

1-dimension: $f(x)=F^{\prime}(x)=\frac{d}{d x} \int_{-\infty}^{x} f(t) d t$.

$$
\text { 2-D: } f(x, y)=\frac{\partial^{2}}{\partial x \partial y} F(x, y)=\frac{\partial^{2}}{\partial x \partial y} \int_{-\infty}^{x} \int_{-\infty}^{y} f(s, t) d t d s .
$$

## Bivariate Cumulative distribution Functions (CDF's)

Definition: The joint (bivariate) cdf of $X$ and $Y$ is

$$
F(x, y) \equiv P(X \leq x, Y \leq y), \text { for all } x, y .
$$

$$
F(x, y)=\left\{\begin{array}{lc}
\sum \sum_{s \leq x, t \leq y} f(s, t) & \text { discrete } \\
\int_{-\infty}^{y} \int_{-\infty}^{x} f(s, t) d s d t & \text { continuous }
\end{array}\right.
$$

## Prooerties

(1) $F(x, y)$ is non-decreasing in both $x$ and $y$.
(1) $\lim _{x \rightarrow-\infty} F(x, y)=\lim _{y \rightarrow-\infty} F(x, y)=0$.
(1) $\lim _{x \rightarrow \infty} F(x, y)=F_{Y}(y)=\operatorname{Pr}(Y \leq y)$
(1) $\lim _{y \rightarrow \infty} F(x, y)=F_{X}(x)=\operatorname{Pr}(X \leq x)$.
(1) $F(x, y)$ is cts from the right in both $x$ and $y$.

## Exercise: Air Conditioner Maintenance (صيانة هكف الهعاء)

A company that services air conditioner units in residences and office blocks is interested in how to schedule its technicians in the most efficient manner

- The random variable $X$, taking the values $1,2,3$ and 4 , is the service time in hours
- The random variable Y, taking the values 1,2 and 3, is the number of air conditioner units

| Y= <br> number <br> of units | X=service time |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  | 1 | 2 | 3 | 4 |
| 1 | 0.12 | 0.08 | 0.07 | 0.05 |
| 2 | 0.08 | 0.15 | 0.21 | 0.13 |
| 13 | 0.01 | 0.01 | 0.02 | 0.07 |

- Joint p.m.f and the total probability
$\sum_{i} \sum_{j} P_{i j}=0.12+0.08+\cdots \cdot+0.07=1$
- Joint cumulative distribution function

$$
\begin{aligned}
F(2,2) & =p_{11}+p_{12}+p_{21}+p_{22} \\
& =0.12+0.18+0.08+0.15 \\
& =0.43
\end{aligned}
$$

## Conditional Probability Distributions (1/3)

- Conditional probability distributions
- The probability of the random variable X under the knowledge provided by the value of Y is given by
- Discrete case

$$
p_{i \mid j}=P(X=i \mid Y=j)=\frac{P(X=i, Y=j)}{P(Y=j)}=\frac{p_{i j}}{p_{+j}}
$$

- Continuous case

$$
f_{X \mid Y=y}(x)=\frac{f(x, y)}{f_{Y}(y)}
$$

- The conditional probability distribution is a probability distribution.

Because a conditional probability mass function $f_{Y \mid x}(y)$ is a probability mass function for all $y$ in $R_{x}$, the following properties are satisfied:
(1) $f_{Y \mid x}(y) \geq 0$
(2) $\sum_{R_{x}} f_{Y \mid x}(y)=1$
(3) $P(Y=y \mid X=x)=f_{Y \mid x}(y)$

## Conditional Probability Distributions (2/3)

## Conditional Mean and Variance

Definition
Let $R_{x}$ denote the set of all points in the range of $(X, Y)$ for which $X=x$. The conditional mean of $Y$ given $X=x$, denoted as $E(Y \mid x)$ or $\mu_{Y \mid x}$, is

$$
\begin{equation*}
E(Y \mid x)=\sum_{R_{x}} y f_{Y \mid x}(y) \tag{5-6}
\end{equation*}
$$

and the conditional variance of $Y$ given $X=x$, denoted as $V(Y \mid x)$ or $\sigma_{Y \mid x}^{2}$, is

$$
V(Y \mid x)=\sum_{R_{x}}\left(y-\mu_{Y \mid x}\right)^{2} f_{Y \mid x}(y)=\sum_{R_{x}} y^{2} f_{Y \mid x}(y)-\mu_{Y \mid x}^{2}
$$

Example (Exercise: Air Conditioner )

- Marginal probability distribution of Y

$$
P(Y=3)=p_{+3}=0.01+0.01+0.02+0.07=0.11
$$

- Conditional distribution of X

$$
p_{1 \mid Y=3}=P(X=1 \mid Y=3)=\frac{p_{13}}{p_{+3}}=\frac{0.01}{0.11}=0.091
$$

## Conditional Probability Distributions (3/3)

$\square$ Example: The marginal probability distribution for X and Y .

| y=number of times <br> city name is stated | $\mathrm{X}=$ number of bars of signal strength |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
| 4 | 1 | 2 | 3 | Marginal probability distribution of Y |
| 3 | 0.15 | 0.1 | 0.05 | 0.3 |
| 2 | 0.02 | 0.1 | 0.05 | 0.17 |
| 1 | 0.02 | 0.03 | 0.2 | 0.25 |
|  | 0.01 | 0.02 | 0.25 | 0.28 |
|  | 0.2 | 0.25 | 0.55 |  |

Marginal probability distribution of $X$

$$
\begin{aligned}
& P(Y=1 \mid X=3)=P(X=3, Y=1) / P(X=3) \\
& \quad=f_{x, y}(3,1) / f_{x}(3)=0.25 / 0.55=0.454 \\
& E(Y \mid 1)
\end{aligned} \quad=\sum_{y} y f_{Y \mid 1}(y) \text {. }
$$

## Expected Values for Jointly Distributed Continuous R.V.s

- Let X and Y be continuous random variables with joint probability density function $\mathrm{f}(\mathrm{x}, \mathrm{y})$. We define $\mathrm{E}(\mathrm{X})$ and $\mathrm{E}(\mathrm{Y})$ as

$$
\mathrm{E}(\mathrm{X})=\int_{-\infty}^{\infty} x \mathrm{f}_{\mathrm{X}}(x) \mathrm{d} x \text { and } \mathrm{E}(\mathrm{Y})=\int_{-\infty}^{\infty} y \mathrm{f}_{\mathrm{Y}}(y) \mathrm{d} y .
$$

- Example. For the random variables X and Y from the previous slide,

$$
\mathrm{f}_{\mathrm{X}}(x)=\mathrm{e}^{-x}, x>0 \text { and } \mathrm{f}_{\mathrm{Y}}(y)=2 \mathrm{e}^{-2 y}, y>0 .
$$

That is, X and Y are exponential random variables. It follows that

$$
\mathrm{E}(\mathrm{X})=1 \text { and } \mathrm{E}(\mathrm{Y})=\frac{1}{2} .
$$

## Independence, Covariance and Correlation (1/13)

## Independence

## Discrete case

Two random variables X and Y are said to be independent if and only if

$$
\mathrm{P}[\mathrm{X}=\mathrm{x}, \mathrm{Y}=\mathrm{y}]=\mathrm{P}[\mathrm{X}=\mathrm{x}] \mathrm{P}[\mathrm{Y}=\mathrm{y}] \quad \text { for all real numbers } \mathrm{x} \text { and } \mathrm{y} \text {. }
$$

This definition of independence for discrete random variables translates into the statement that X and Y are independent if and only if a cell value is the product of the row total times the column total. i.e

$$
p_{i j}=p_{i+} p_{+j} \text { for all values } i \text { of } X \text { and } j \text { of } Y
$$

## Continuous case

For continuous random variables, the condition for independence of X and Y becomes $X$ and $Y$ are independent if and only if

$$
f(x, y)=f_{X}(x) f_{Y}(y) \text { for all } x \text { and } y
$$

The marginal density functions can be multiplied together to produce the joint density function. Thus the random variables X and Y are independent.

## Independence, Covariance and Correlation (2/13)

## Example: (Discrete case)

Are the random variables X and Y described above with the following joint probability density table independent?

| Y Values |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| X Values |  | 0 | 1 | 2 | 3 |  |
|  | 0 | 1/8 | 0 | 0 | 0 | 1/8 |
|  | 1 | 0 | 1/8 | 1/8 | 1/8 | 3/8 |
|  | 2 | 0 | 1/4 | 1/8 | 0 | 3/8 |
|  | 3 | 0 | 1/8 | 0 | 0 | 1/8 |
|  |  | 1/8 | 1/2 | 1/4 | 1/8 |  |

The random variables are not independent because, for example $\mathrm{P}[\mathrm{X}=0, \mathrm{Y}=1]=0$ but $\mathrm{P}[\mathrm{X}=0]=1 / 8$ and $\mathrm{P}[\mathrm{Y}=1]=4 / 8$.

## Example: (continuous case)

For the joint density function $\mathrm{f}(\mathrm{x}, \mathrm{y})=1$ for x on $[0,1]$ and y on $[0,1]$ and 0 otherwise, the marginal density function of $X, f_{X}(x)=1$ for $x$ on $[0,1]$ and the marginal density function of $\mathrm{Y}, \mathrm{f}_{\mathrm{Y}}(\mathrm{y})=1$ for y on $[0,1]$. The marginal density functions can be multiplied together to produce the joint density function. Thus the random variables X and Y are independent

## Independence, Covariance and Correlation (3/13)

## Functions of Independent Random Variables

Theorem. Let X and Y be independent random variables and let $g$ and $h$ be real valued functions of a single real variable.

Then
(i) $g(X)$ and $h(Y)$ are also independent random variables
(ii) $\mathrm{E}[\mathrm{g}(\mathrm{X}) \mathrm{h}(\mathrm{Y})]=\mathrm{E}[\mathrm{g}(\mathrm{X})] \mathrm{E}[\mathrm{h}(\mathrm{Y})]$.

- Example. If X and Y are independent, then $\mathrm{E}\left[(\sin \mathrm{X}) \mathrm{e}^{\mathrm{Y}}\right]=\mathrm{E}[\sin \mathrm{X}] \mathrm{E}\left[\mathrm{e}^{\mathrm{Y}}\right]$.


## Independence, Covariance and Correlation (4/13)

## Covariance

$$
\begin{aligned}
\operatorname{Cov}(X, Y) & =E((X-E(X))(Y-E(Y))) \\
& =E(X Y)-E(X) E(Y) \\
\operatorname{Cov}(X, Y) & =E((X-E(X))(Y-E(Y))) \\
& =E(X Y-X E(Y)-E(X) Y+E(X) E(Y)) \\
& =E(X Y)-E(X) E(Y)-E(X) E(Y)+E(X) E(Y) \\
& =E(X Y)-E(X) E(Y)
\end{aligned}
$$

- May take any positive or negative numbers.
- Independent random variables have a covariance of zero
- What if the covariance is zero?
$\mathrm{E}(\mathrm{XY})=\mathrm{E}(\mathrm{X}) \mathrm{E}(\mathrm{Y})$


## Independence, Covariance and Correlation (5/13)

## General Properties of $(\operatorname{Cov}(\mathbf{X}, \mathbf{Y}))$

$$
\begin{gathered}
4-\operatorname{Cov}\left(\sum_{i=1}^{n} X_{i}, \sum_{j=1}^{m} Y_{j}\right)=\sum_{i=1}^{n} \sum_{j=1}^{m} \operatorname{Cov}\left(X_{i}, Y_{j}\right) \\
5-\operatorname{Var}\left(\sum_{i=1}^{n} X_{i}\right)=\operatorname{Cov}\left(\sum_{i=1}^{n} X_{i}, \sum_{j=1}^{n} X_{i}\right) \\
=\sum_{i=1}^{n} \operatorname{Var}\left(X_{i}\right)+\sum_{i=j} \sum \operatorname{Cov}\left(X_{i}, X_{j}\right) \\
=\sum_{i=1}^{n} \operatorname{Var}\left(X_{i}\right)+2 \sum_{i<j} \sum \operatorname{Cov}\left(X_{i}, X_{j}\right)
\end{gathered}
$$

## Independence, Covariance and Correlation (6/13)

- Properties of $\operatorname{Cov}\left(X_{1}, X_{2}\right)$
$\operatorname{Cov}\left(X_{1}, X_{2}\right)=E\left[X_{1} X_{2}\right]-E\left[X_{1}\right] E\left[X_{2}\right]$
, $\operatorname{Cov}(\mathbf{a x}, \mathrm{bY})=\mathbf{a b} \operatorname{Cov}(\mathbf{X}, \mathbf{Y})$
$\Rightarrow \operatorname{Cov}\left(X_{1}, X_{2}\right)=\operatorname{Cov}\left(X_{2}, X_{1}\right)$
$\operatorname{Cov}\left(X_{1}, X_{1}\right)=\operatorname{Var}\left(X_{1}\right) \quad \operatorname{Cov}\left(X_{2}, X_{2}\right)=\operatorname{Var}\left(X_{2}\right)$
$\operatorname{Cov}\left(X_{1}+X_{2}, X_{1}\right)=\operatorname{Cov}\left(X_{1}, X_{1}\right)+\operatorname{Cov}\left(X_{2}, X_{1}\right)$
$\operatorname{Cov}\left(X_{1}+X_{2}, X_{1}+X_{2}\right)=\operatorname{Cov}\left(X_{1}, X_{1}\right)+\operatorname{Cov}\left(X_{1}, X_{2}\right)+$ $\operatorname{Cov}\left(X_{2}, X_{1}\right)+\operatorname{Cov}\left(X_{2}, X_{2}\right)$
$\Rightarrow \operatorname{Var}\left(X_{1}+X_{2}\right)=\operatorname{Var}\left(X_{1}\right)+\operatorname{Var}\left(X_{2}\right)+2 \operatorname{Cov}\left(X_{1}, X_{2}\right)$


## Independence, Covariance and Correlation (7/13)

## General Properties of $(\operatorname{Cov}(\mathbf{X}, \mathbf{Y}))$

$$
\begin{aligned}
4 & -\operatorname{Cov}\left(\sum_{i=1}^{n} X_{i}, \sum_{j=1}^{m} Y_{j}\right)=\sum_{i=1}^{n} \sum_{j=1}^{m} \operatorname{Cov}\left(X_{i}, Y_{j}\right) \\
5 & -\operatorname{Var}\left(\sum_{i=1}^{n} X_{i}\right)=\operatorname{Cov}\left(\sum_{i=1}^{n} X_{i}, \sum_{j=1}^{n} X_{i}\right) \\
& =\sum_{i=1}^{n} \operatorname{Var}\left(X_{i}\right)+\sum_{i \neq j} \sum \operatorname{Cov}\left(X_{i}, X_{j}\right) \\
& =\sum_{i=1}^{n} \operatorname{Var}\left(X_{i}\right)+2 \sum_{i<j} \sum \operatorname{Cov}\left(X_{i}, X_{j}\right)
\end{aligned}
$$

## Independence, Covariance and Correlation (8/13)

- Example (Air conditioner maintenance)

$$
\begin{aligned}
& E(X)=2.59, \quad E(Y)=1.79 \\
& E(X Y)=\sum_{i=1}^{4} \sum_{j=1}^{3} i j p_{i j} \\
& =(1 \times 1 \times 0.12)+(1 \times 2 \times 0.08) \\
& +\cdots+(4 \times 3 \times 0.07)=4.86 \\
& \operatorname{Cov}(X, Y)=E(X Y)-E(X) E(Y) \\
& =4.86-(2.59 \times 1.79)=0.224
\end{aligned}
$$

Exercise: Find $\operatorname{Var}(\mathbf{X})$, $\operatorname{Var}(\mathbf{Y})$

## Independence, Covariance and Correlation (9/13)

## Correlation

- The correlation is a measure of the linear relationship between X and Y . It is obtained by dividing the covariance by the product of the two standard deviations, i.e.,

$$
\operatorname{Corr}(X, Y)=\frac{\operatorname{Cov}(X, Y)}{\sqrt{\operatorname{Var}(X) \operatorname{Var}(Y)}}
$$

- The correlation assumes values between -1 and 1 .
- A value close to 1 implies a strong positive relationship.
- A value close to -1 implies a strong negative relationship.
- A value close to zero implies little or no relationship.
- The independent random variables have a correlation of zero.


## Independence, Covariance and Correlation (10/13)

An important implication of independence
Suppose that the components X and Y of the discrete bivariate random variable $(\mathrm{X}, \mathrm{Y})$ are independent. Then its covariance ic zern

Always

$$
\begin{aligned}
\sigma_{\mathrm{XY}} & =\operatorname{cov}(\mathrm{X}, \mathrm{Y}) \\
& =\sum_{\mathrm{x} \in \mathrm{X}(\Omega)} \sum_{\mathrm{y} \in \mathrm{Y}(\Omega)}\left(\mathrm{x}-\mu_{\mathrm{X}}\right)\left(\mathrm{y}-\mu_{\mathrm{Y}}\right) \mathrm{P}(\mathrm{X}=\mathrm{x} \wedge \mathrm{Y}=\mathrm{y}) \\
& \left.=\sum_{\mathrm{x} \in \mathrm{X}(\Omega)} \sum_{\mathrm{y} \in \mathrm{Y}(\Omega)}\left(\mathrm{x}-\mu_{\mathrm{X}}\right)\left(\mathrm{y}-\mu_{\mathrm{Y}}\right) \mathrm{P}(\mathrm{X}=\mathrm{x}) \mathrm{T}=\mathrm{y}\right) \\
& =\sum_{\mathrm{x} \in \mathrm{X}(\Omega)}\left(\mathrm{x}-\mu_{\mathrm{X}}\right) \mathrm{P}(\mathrm{X}=\mathrm{x})\left[\sum_{\mathrm{y} \in \mathrm{Y}}^{\Sigma} /(\Omega)\right. \\
& =\mathrm{E}\left(\mathrm{H}-\mu_{\mathrm{Y}}\right) \mathrm{E}\left(\mathrm{Y}-\mu_{\mathrm{Y}}\right) \\
& =0
\end{aligned}
$$

## Independence, Covariance and Correlation (11/13)

## The continuous case

$$
\begin{aligned}
\sigma_{\mathrm{XY}} & =\int_{-\infty}^{\infty} \int_{-\infty}^{\infty}\left(\mathrm{x}-\mu_{\mathrm{X}}\right)\left(\mathrm{y}-\mu_{\mathrm{Y}}\right) \mathrm{f}_{\mathrm{XY}}(\mathrm{x}, \mathrm{y}) \mathrm{dydx} \\
& =\int_{-\infty}^{\infty} \int_{-\infty}^{\infty}\left(\mathrm{x}-\mu_{\mathrm{X}}\right)\left(\mathrm{y}-\mu_{\mathrm{Y}}\right) \mathrm{f}(\mathrm{x}) \mathrm{f}(\mathrm{y}) \mathrm{dydx} \\
& =\int_{-\infty}^{\infty}\left(\mathrm{x}-\mu_{\mathrm{x}}\right) \mathrm{f}(\mathrm{x}) \int_{-\infty}^{\infty}\left(\mathrm{y}-\mu_{\mathrm{Y}}\right) \mathrm{f}(\mathrm{y}) \mathrm{dydx} \\
& =\int_{-\infty}^{\infty}\left(\mathrm{x}-\mu_{\mathrm{X}}\right) \mathrm{f}(\mathrm{x})\left[\int_{-\infty}^{\infty}\left(\mathrm{y}-\mu_{\mathrm{Y}}\right) \mathrm{f}(\mathrm{y}) \mathrm{dy}\right] \mathrm{dx} \\
& =\left[\int_{-\infty}^{\infty}\left(\mathrm{y}-\mu_{\mathrm{Y}}\right) \mathrm{f}(\mathrm{y}) \mathrm{dy}\right]_{-\infty}^{\infty}\left(\mathrm{x}-\mu_{\mathrm{X}}\right) \mathrm{f}(\mathrm{x}) \mathrm{dx} \\
& =\mathrm{E}\left(\mathrm{Y}-\mu_{\mathrm{Y}}\right) \mathrm{E}\left(\mathrm{X}-\mu_{\mathrm{x}}\right) \\
& =0 .
\end{aligned}
$$

## Independence, Covariance and Correlation (12/13)

- Example : (Air conditioner maintenance)

$$
\begin{aligned}
& \operatorname{Var}(X)=1.162, \quad \operatorname{Var}(Y)=0.384 \\
& \begin{aligned}
& \operatorname{Corr}(X, Y)= \\
& \sqrt{\sqrt{\operatorname{Var}(X) \operatorname{Var}(Y)}} \\
&=\frac{0.224}{\sqrt{1.162 \times 0.384}}=0.34
\end{aligned}
\end{aligned}
$$

## Independence, Covariance and Correlation (13/13)

Moments of bivariate random variables

Joint moments of (X,Y):

```
EX, EY
EX }\mp@subsup{}{}{2},\textrm{EXY},\mp@subsup{E}{}{2
```



Joint central moments of (X,Y):
E(X-EX), E(Y-EY)
$\mathrm{E}(\mathrm{X}-\mathrm{EX})^{2}, \mathrm{E}(\mathrm{X}-\mathrm{EX})(\mathrm{Y}-\mathrm{EY}), \mathrm{E}(\mathrm{Y}-\mathrm{EY})^{2}$
$E(X-E X)^{3}, E(X-E X)^{2}(Y-E Y), E(X-E X)(Y-E Y)^{2}, E(Y-E Y)^{3}$

- I read from my textbook that $\operatorname{cov}(\mathrm{X}, \mathrm{Y})=0$ does not guarantee X and Y are independent. But if they are independent, their covariance must be 0 . I could not think of any proper example yet; could someone provide one?
- Easy example: Let X be a random variable that is -1 or +1 with probability 0.5 . Then let $Y$ be a random variable such that $\mathrm{Y}=0$ if $\mathrm{X}=-1$, and Y is randomly -1 or +1 with probability 0.5 if $\mathrm{X}=1$.
- Clearly X and Y are highly dependent (since knowingY allows me to perfectly know X ), but their covariance is zero: They both have zero mean, and

$$
\begin{aligned}
\mathbb{E}[X Y] & =(-1) \cdot 0 \quad \cdot P(X=-1) \\
& +1 \quad \cdot 1 \quad \cdot P(X=1, Y=1) \\
& +1 \quad \cdot(-1) \cdot P(X=1, Y=-1) \\
& =0 .
\end{aligned}
$$

- Or more generally, take any distribution $\mathrm{P}(\mathrm{X})$ and any $\mathrm{P}(\mathrm{Y} \mid \mathrm{X})$ such that $P(Y=a \mid X)=P(Y=-a \mid X)$ for all $X$ (i.e., a joint distribution that is symmetric around the x axis), and you will always have zero covariance. But you will have non-independence whenever $\mathrm{P}(\mathrm{Y} \mid \mathrm{X}) \neq \mathrm{P}(\mathrm{Y})$; i.e., the conditionals are not all equal to the marginal. Or ditto for symmetry around the y axis.


## Exercices

(1) Let the joint distribution of $X$ and $Y$ is given by :

$$
f(x, y)=c x y, x=1,2,3 ; y=1,2,3 .
$$

Find
(i) the constant c ,
(ii) (ii) $\mathrm{p}(\mathrm{x}=2, \mathrm{y}=3)$,
(iii) $\mathrm{P}(1 \leq \mathrm{X} \leq 2, \mathrm{Y} \leq 2) \mathrm{P}(\mathrm{Y}<2), \mathrm{p}(\mathrm{X}=1) \mathrm{p}(\mathrm{Y}=3)$

Find the marginal probability mass functions of X and Y above and determine whether x and Y are independent
(2) Let $X, Y$ have joint density functions

$$
f(x, y)=\left\{\begin{array}{l}
c\left(x^{2}+y^{2}\right) \quad 0 \leq x \leq 1,0 \leq y \leq 1 \\
0, \text { otherwise }
\end{array}\right.
$$

## Determine

(a) the constant c
(b) $\mathrm{P}(\mathrm{X}<1 / 2, \mathrm{Y}>1 / 2), \mathrm{P}(1 / 4<\mathrm{X}<3 / 4)$
(d) $\mathrm{P}(\mathrm{Y}<1 / 4)$
(e) Find the marginal distribution functions of X and Y .
32) Are X, Y independent
(3) The joint distributions are given in each question. Find the conditional distribution of
(a) X given Y
(b) Y given X

$$
\begin{aligned}
\text { 2.58) } f(x, y) & =\left(\frac{y y}{36}\right), \\
x & =1,2,3 \text { and } y=1,2,3
\end{aligned}
$$

(2.59) $f(x, y)= \begin{cases}\frac{3}{2}\left(x^{2}+y^{2}\right), & 0 \leq x \leq 1,0 \leq y \leq 1 \\ 0 & , \text { otherwise }\end{cases}$
(2.60) $\mathrm{f}(\mathrm{x}, \mathrm{y})= \begin{cases}(\mathrm{x}+\mathrm{y}), & 0 \leq x \leq 1,0 \leq y \leq 1 \\ 0 & , \text {,therwise }\end{cases}$

$$
\text { (2.61) } \mathrm{f}(\mathrm{x}, \mathrm{y})= \begin{cases}e^{-(\mathrm{x}+\mathrm{y})}, & x \geq 0, y \geq 0 \\ 0 & \text {, otherwise }\end{cases}
$$

