Introduction to Probability and Statistics

Prof. Mohammed-Slim Alouini & Prof. Hesham ElSawy
slim.alouini@kaust.edu.sa & hesham.elsawy@kfupm.edu.sa

Eng. Chaouki Bem Issaid & Eng. Lama Niyazi
chaouki.benissaid@kaust.edu.sa & lama.niyazi@kaust.edu.sa

STC Academy, Riyadh, KSA

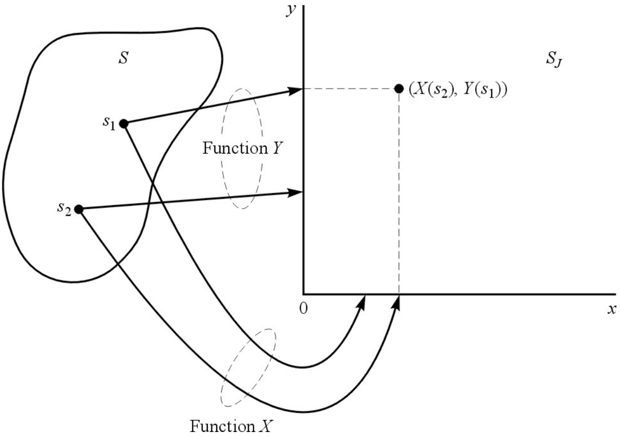
23 June to 4 July 2019

Lecture Objectives

- Progress
  - Basic concepts
  - Discrete random variable
  - Continuous random variable
  - Generation of random variables and Inequalities
- Today
  - Two random variables

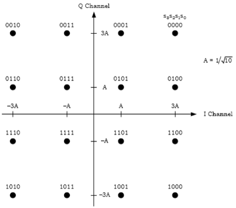
Two Random Variables

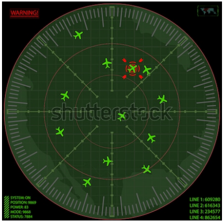
- Consider two random variables X and Y that are defined on an outcome space S

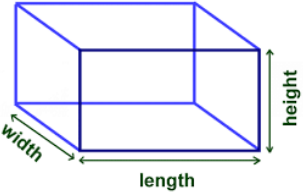


Example

- Digitally modulated signals
- Two-dimensional location on a radar screen
- Dimensions of objects over a production belt







Example

- Consider that 3 balls will be selected at random from an urn with 3 red balls, 4 black balls, and 5 blue balls. Let X and Y be the number of selected red and black balls respectively
- The outcome space is $\{{\color{red}\bullet}\,{\color{red}\bullet}\,{\color{red}\bullet}\,\bullet\,\bullet\,\bullet\,\bullet\,{\color{blue}\bullet}\,{\color{blue}\bullet}\,{\color{blue}\bullet}\,{\color{blue}\bullet}\,{\color{blue}\bullet}\}$
- The events are $\{{\color{red}\bullet}{\color{red}\bullet}{\color{red}\bullet}\}$ $\{{\color{red}\bullet}{\color{red}\bullet}{\color{blue}\bullet}\}$ $\{{\color{red}\bullet}{\color{blue}\bullet}{\color{red}\bullet}\}$ $\{{\color{blue}\bullet}{\color{red}\bullet}{\color{red}\bullet}\}$ $\{{\color{red}\bullet}{\color{blue}\bullet}{\color{blue}\bullet}\}$ $\{{\color{blue}\bullet}{\color{red}\bullet}{\color{blue}\bullet}\}$ $\{{\color{blue}\bullet}{\color{blue}\bullet}{\color{red}\bullet}\}$ $\{{\color{red}\bullet}{\color{red}\bullet}\bullet\}$ $\{{\color{red}\bullet}\bullet{\color{red}\bullet}\}$ $\{\bullet{\color{red}\bullet}{\color{red}\bullet}\}$ $\{{\color{red}\bullet}\bullet\bullet\}$ $\{\bullet{\color{red}\bullet}\bullet\}$ $\{\bullet\bullet{\color{red}\bullet}\}$ $\{{\color{blue}\bullet}{\color{blue}\bullet}{\color{blue}\bullet}\}$ $\{{\color{blue}\bullet}{\color{blue}\bullet}\bullet\}$ $\{{\color{blue}\bullet}\bullet{\color{blue}\bullet}\}$ $\{\bullet{\color{blue}\bullet}{\color{blue}\bullet}\}$ $\{{\color{blue}\bullet}\bullet\bullet\}$ $\{\bullet{\color{blue}\bullet}\bullet\}$ $\{\bullet\bullet{\color{blue}\bullet}\}$ $\{\bullet\bullet\bullet\}$ $\ldots$
- The possible random variables realizations are $(3,0)$ $(2,0)$ $(1,0)$ $(0,0)$ $(0,3)$ $(0,2)$ $(0,1)$ $(2,1)$ $(1,2)$ $(1,1)$
- The bivariate distribution of X and Y is

$i \backslash j$	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 $= P\{Y = j\}$	$\frac{56}{220}$	$\frac{112}{220}$	$\frac{48}{220}$	$\frac{4}{220}$	

Bivariate Discrete Distribution

- The joint probability mass function is given by

$$P_{X,Y}(x,y) = \mathbb{P}(X = i, Y = j) = \mathbb{P}(X = i \text{ and } Y = j)$$

- The range of X and Y is defined as

$$R_{X,Y}(x,y) = \{(x,y)|\mathbb{P}(X = i, Y = j) > 0\}$$

- The range $R_{X,Y} \subset R_X \times R_Y$

$$R_{X,Y}(x,y) \subset \{(x_i, y_j)|x_i \in R_X, y_j \in R_Y\}$$

- If we defined the range $R_{X,Y} = R_X \times R_Y$, we should keep in mind that $P_{XY}(x,y)$ can be zero for some pairs (x,y)

Bivariate PMF

- For discrete random variable

$$\sum_{(x_i,y_i)\in R_{XY}} P_{X,Y}(x,y) = 1$$

- The probability of an event $\mathcal{A} \subset \mathbb{R}^2$ is calculated as

$$\mathbb{R}((X,Y) \in \mathcal{A}) = \sum_{(x_i,y_i)\in(\mathcal{A}\cap R_{XY})} \mathbb{P}(X = i, Y = j)$$

- The marginal probability mass functions of X and Y are

$$P_X(x) = \sum_{y_j \in R_Y} P_{X,Y}(x_i, y_j) \quad \text{and} \quad P_Y(y) = \sum_{x_i \in R_X} P_{X,Y}(x_i, y_j)$$

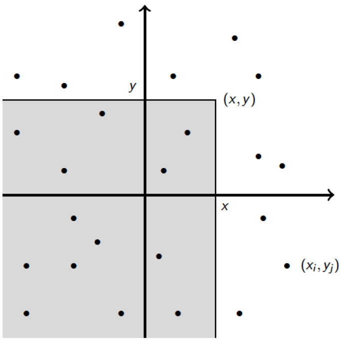
- The conditional probability mass function of X at $Y = j$ and Y at $X = j$ are

$$P_X(x|Y = j) = \frac{P_{X,Y}(x,j)}{P_Y(j)} \quad \text{and} \quad P_Y(y|X = i) = \frac{P_{X,Y}(i,Y)}{P_X(i)}$$

The CDF

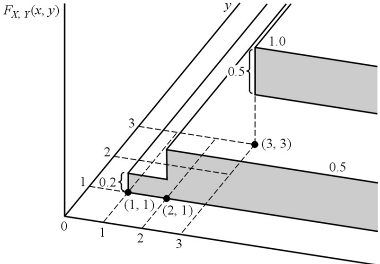
- The joint probability distribution function is given by

$$F_{X,Y}(x, y) = \sum_{x_i < x} \sum_{y_i < y} \mathbb{P}(X = x_i, Y = y_j)$$

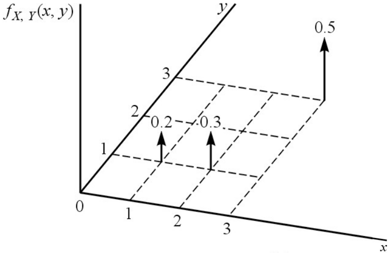


Example

- Consider a bivariate discrete distribution that has elements at $(1,1)$, (2,1), and (3,3), where $\mathbb{P}(1,1) = 0.2$, $\mathbb{P}(2,1) = 0.3$, and $\mathbb{P}(3,3) = 0.5$



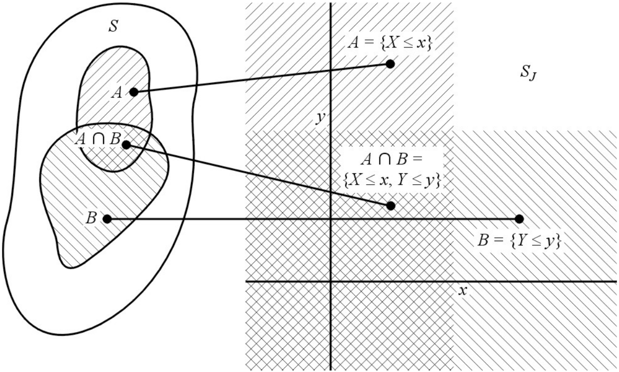
(a)



(b)

Bivariate Random Variable

- Define the events $\mathcal{A} = \{X \leq x\}$ and $\mathcal{B} = \{Y \leq y\}$



Joint Distribution Functions

- For the bivariate case, the joint distribution function is defined as

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

Properties of Distribution Functions

- The CDF has the following properties

1. $F_{X,Y}(-\infty, -\infty) = 0$, $F_{X,Y}(x, -\infty) = 0$ $F_{X,Y}(-\infty, y) = 0$
2. $F_{X,Y}(\infty, \infty) = 1$
3. $0 \leq F_{X,Y}(x, y) \leq 1$
4. $F_{X,Y}(x, y)$ is a nondecreasing function is both x and y
5. $\mathbb{P}(x_1 < X \leq x_2, y_1 < Y \leq y_2) =$ $F_{X,Y}(x_2, y_2) - F_{X,Y}(x_1, y_2) - F_{X,Y}(x_2, y_1) + F_{X,Y}(x_1, y_1)$
6. $F_{X,Y}(x, \infty) = F_X(x)$ & $F_{X,Y}(\infty, y) = F_Y(y)$

Example

- The joint distribution is

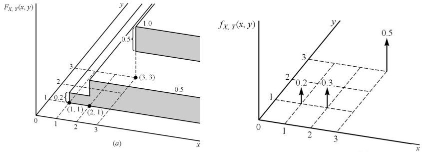
$$F_{X,Y}(x,y) = 0.2u(x-1)u(y-1) + 0.3u(x-2)u(y-1) + 0.5u(x-3)u(y-3)$$

- The marginal distribution of X

$$\begin{aligned} F_X(x) &= F_{X,Y}(x,\infty) \\ &= 0.2u(x-1) + 0.3u(x-2) + 0.5u(x-3) \end{aligned}$$

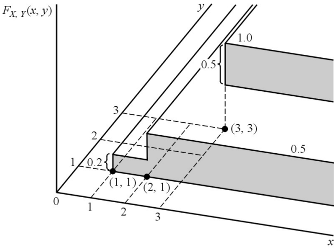
- The marginal distribution of Y

$$\begin{aligned} F_Y(y) &= F_{X,Y}(\infty,y) \\ &= 0.2u(y-1) + 0.3u(y-1) + 0.5u(y-3) \\ &= 0.5u(y-1) + 0.5u(y-3) \end{aligned}$$



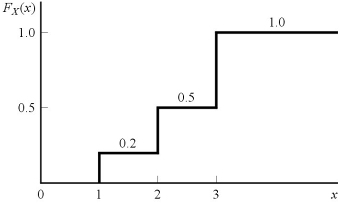
(a) (b)

Example



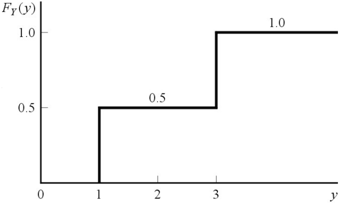
$$F_{X,Y}(x,y) = 0.2u(x-1)u(y-1) + 0.3u(x-2)u(y-1) + 0.5u(x-3)u(y-3)$$

$$F_X(x) = 0.2u(x-1) + 0.3u(x-2) + 0.5u(x-3)$$



(a)

$$F_Y(y) = 0.5u(y-1) + 0.5u(x-3)$$



(b)

Joint Density Functions

- For the bivariate case, the joint densify function is defined as

$$P_{X,Y}(x,y) = \mathbb{P}(X = x_i, Y = y_j)$$

Properties of Density Functions

- The PDF has the following properties
  1. $P_{X,Y}(x,y) > 0$
  2. $\sum_{(x,y)\in R_{XY}} PX, Y(x,y) = 1$
  3. $F_{X,Y}(x,y) = \sum_{x_i \le x} \sum_{y_j \le y} P_{X,Y}(x_i, y_j)$
  4. $F_X(x) = \sum_{x_i \le x} \sum_{y_j \in R_y} P_{X,Y}(x_i, y_j)$
     $F_Y(y) = \sum_{x_i \in R_X} \sum_{y_j \le y} P_{X,Y}(x_i, y_j)$
  5. $\mathbb{P}(x_1 \le X \le x_2, y_1 \le Y \le y_2) = \sum_{x_1 \le x_i \le x_2} \sum_{y_1 \le y_j \le y_2} P_{X,Y}(x,y)$
  6. $P_X(x_i) = \sum_{y_j \in R_Y} P_{X,Y}(x_i, y_j)$
     $P_Y(y_j) = \sum_{x_i \in R_X} P_{X,Y}(x_i, y_j)$

- Caution: Sometimes the summation boundaries are dependent

Conditional Density & Distribution Functions

- The conditional distribution function is defined as

$$F_{X|B}(x|B) = \mathbb{P}(X \leq x|B) = \frac{\mathbb{P}(X \leq x \cap B)}{\mathbb{P}(B)}$$

- The conditional density function is obtained as

$$P_{X|B}(x_i|B) = \frac{\mathbb{P}\{X = x_i, B\}}{\mathbb{P}(B)}$$

- The event B can be defined as
    - Point conditioning
    - Interval conditioning

Point Conditioning

Conditional density functions: point conditioning

- The event B is defined as $\mathcal{B} = Y = y_i$
- The conditional CDF is

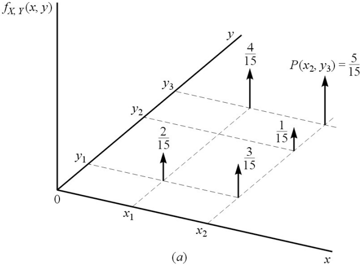
$$F_{X|Y}(x|Y = y_k) = \sum_{x_i < x} \frac{\mathbb{P}(x_i, y_k)}{\mathbb{P}(y_k)}$$

- The conditional PMF is

$$P_{X|Y}(x_i|Y = y_k) = \frac{\mathbb{P}(x_i, y_k)}{\mathbb{P}(y_k)}$$

Example

- Joint and conditional PDFs of a discrete random variable



Interval Conditioning

Conditional density functions: interval conditioning

- The event B is defined as $B = \{y_a \leq Y \leq y_b\}$
- The conditional density function is

$$F_{X|Y}(x|y_a \leq Y \leq y_b) = \frac{\sum_{x_i < x} \sum_{y_j = y_a}^{y_b} P_{X,Y}(x_i, y_j)}{\sum_{y_j = y_a}^{y_b} P_Y(y_j)}$$

- For discrete random variable

$$F_{X|Y}(x|y_a \leq Y \leq y_b) = \frac{F_{X,Y}(x, y_b) - F_{X,Y}(x, y_a)}{F_Y(y_b) - F_Y(y_a)}$$

$$P_{X|Y}(x_i|y_a \leq Y \leq y_b) = \frac{\sum_{y_j = y_a}^{y_b} P_{X,Y}(x_i, y_j)}{\sum_{y_j = y_a}^{y_b} P_Y(y_j)}$$

Statistical Independence

- Two events A and B are statistically independent if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

- The statistical independence for random variables can be stablished be defining $A = \{X \leq x\}$ and $B = \{Y \leq y\}$

Statistical Independence

- Two random variables X and Y are statistically independent iff

$$F_{X,Y}(x, y) = F_X(x)F_Y(y)$$

- Equivalently,

$$P_{X,Y}(x_i, y_j) = P_X(x_i)P_Y(y_j)$$

- Hence,

$$P_{X|Y}(x_i|y_j) = P_X(x_i) \quad \text{and} \quad P_{Y|X}(y_j|x_i) = P_Y(y_j)$$

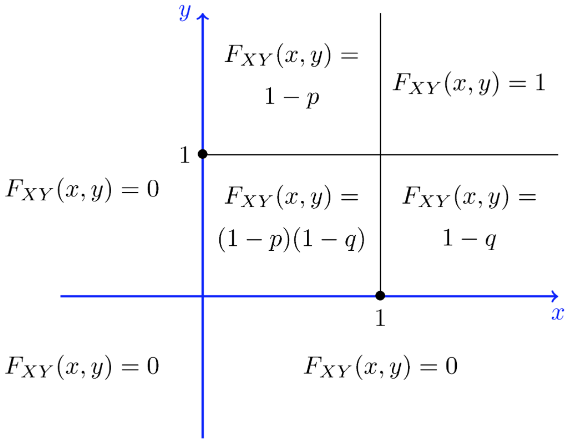
Example

- Let $X \sim Bernoulli(p)$ and $Y \sim Bernoulli(q)$ be independent random variables, where $0 \leq p, q \leq 1$.
- Find the joint PMF and CDF of X and Y
- The range $R_{XY} = \{(0,0), (0,1), (1,0), (1,1)\}$

$$P_{XY}(x_i, y_j) = \begin{cases} (1-p)(1-q) & \text{for } x_i = 0 \text{ and } y_j = 0 \\ (1-p)q & \text{for } x_i = 0 \text{ and } y_j = 1 \\ p(1-q) & \text{for } x_i = 1 \text{ and } y_j = 0 \\ pq & \text{for } x_i = 1 \text{ and } y_j = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$F_{XY}(x, y) = \begin{cases} 0 & \text{for } x < 0 \text{ or } y < 0 \\ (1-p)(1-q) & \text{for } 0 \leq x < 1 \text{ and } 0 \leq y < 1 \\ (1-p) & \text{for } 0 \leq x < 1 \text{ and } y > 1 \\ (1-q) & \text{for } x > 1 \text{ and } 0 \leq y < 1 \\ 1 & \text{for } x > 1 \text{ and } y > 0 \end{cases}$$

Example



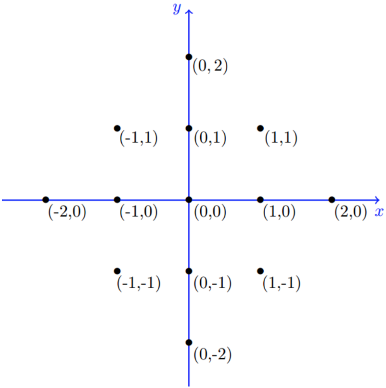
y
$F_{XY}(x,y) =$
$1-p$
$F_{XY}(x,y) = 1$
1
$F_{XY}(x,y) = 0$
$F_{XY}(x,y) =$
$(1-p)(1-q)$
$F_{XY}(x,y) =$
$1-q$
1
x
$F_{XY}(x,y) = 0$
$F_{XY}(x,y) = 0$

Example

- Consider the set of equiprobable points in the grid shown in the below, with the range G defined as

$$G = \{(x, y)|x, y \in \mathbb{Z}, |x| + |y| \leq 2\}$$

- Find
  - the joint and marginal PMFs of X and Y
  - the conditional PMF of X given Y = 1
- Are X and Y independent?



Conditional Expectation

- Conditional expectation is similar to the ordinary single random variable expectation but using the conditional PMF

Conditional Expectation

- Two X and Y be two discrete random variable then,

$$\mathbb{E}[X|Y=y_i] = \sum_{x_i \in R_X} x_i P_{X|Y}(x_i|y_i)$$

Law of Total Probability

Law of Total Probability

- Since the outcomes of X and Y are disjoint, we have

$$P_X(x) = \sum_{y_i \in R_R} P_{X,Y}(x_i, y_i) = \sum_{y_i \in R_Y} P_{X|Y}(x_i|y_i) P_Y(y_i)$$

- For any set A

$$P_X(x \in A) = \sum_{y_i \in R_Y} P_{X|Y}(X \in A | Y = y_i) P_Y(y_i)$$

- Applying the same concept to expectation

$$\mathbb{E}[X] = \sum_{y_i \in R_Y} \mathbb{E}[X|Y = y_i] P_Y(y_i) = \mathbb{E}[\mathbb{E}[X|Y]]$$

- This is denoted as the law of iterated expectation

Example

- Suppose that the number of customers visiting STC branch in a given day is $N \sim Poisson(\lambda)$. Assume that each customer purchases a sim card with probability p, independently from other customers and independently from the value of N. Let X be the number of customers who purchase sim cards.
- Find $\mathbb{E}[X]$.
- Consider each customer as a trial with success if he purchases a sim card and failure otherwise
- Hence, each customer represent an independent Bernoulli trial
- For a given $N = n$, the number of customers that purchase sim cards have a Binomial distribution with mean np. Hence,

$$\mathbb{E}[X|N = n] = np$$

- Using the law of total probability for expectations, we have

$$\mathbb{E}[X] = \sum_{n_i=0}^{\infty} \mathbb{E}[X|N = n_i]P_N(n_i) = \sum_{n_i=0}^{\infty} n_i p P_N(n_i) = p\mathbb{E}[N] = p\lambda$$

Progress

- Last section
  - Multivariate random variables
  - Joint, marginal, and conditional CDF
  - Joint, marginal, and conditional PMF
  - Statistical independence
- Current section
  - Expectation and Variance

Expected value

Expected value of a function of random variables

- Let $g(X,Y)$ be a function that involve two discrete random variables X and Y. Then, the mean value of $g(X,Y)$ is calculated as follows

$$\mathbb{E}[g(X,Y)] = \sum_{x_i \in R_X} \sum_{y_i \in R_Y} g(x_i, y_j) \mathbb{P}_{X,Y}(x_i, y_j)$$

- For $g(X,Y) = X^n Y^k$, we get the $(n+k)$-order moment about the origin

$$\mathbb{E}[X^n Y^k] = \sum_{x_i \in R_X} \sum_{y_i \in R_Y} x_i^n \, y_j^k \, \mathbb{P}(x_i, y_j)$$

- For $g(X,Y) = (X - \bar{X})^n (Y - \bar{Y})^k$, we get the joint $(n+k)$-order moment about the mean

$$\mathbb{E}[(X - \bar{X})^n (Y - \bar{Y})^k] = \sum_{x_i \in R_X} \sum_{y_i \in R_Y} (x_i - \bar{X})^n (y_j - \bar{Y})^k \, \mathbb{P}(x_i, y_j)$$

Moments about the origin

- First order moments about the origin

$$\mathbb{E}[X] \quad \text{and} \quad \mathbb{E}[Y]$$

- Second order moments about the origin

$$\mathbb{E}[Y^2] \quad \text{and} \quad \mathbb{E}[XY]$$

- Third order moments about the origin

$$\mathbb{E}[X^3] \quad \text{and} \quad \mathbb{E}[Y^3]$$
$$\mathbb{E}[X^2Y] \quad \text{and} \quad \mathbb{E}[XY^2]$$

Correlation

Correlation

The second moment about the origin $\mathbb{E}[\mathbb{XY}]$ is of special interest and is denoted as the correlation of X and Y

$$\mathbb{E}[XY] = \sum_{x_i \in R_X} \sum_{y_i \in R_Y} x_i \, y_j \, \mathbb{P}(x_i, y_j)$$

- $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$, then X and Y are ***uncorrelated***
- Uncorrelation does not imply independence
- Independence implies uncorrelation
- If $\mathbb{E}[XY] = 0$, then X and Y are ***orthogonal***

Joint Central Moments

- First order central moments about the origin are zeros
- Second order moments about the origin (variance and covariance)

$$\mathbb{E}[(X-\bar{X})^2] \quad \text{and} \quad \mathbb{E}[(Y-\bar{Y})^2] \quad \text{and} \quad \mathbb{E}[(X-\bar{X})(Y-\bar{Y})]$$

- Third order moments about the origin

$$\mathbb{E}[(X-\bar{X})^3] \quad \text{and} \quad \mathbb{E}[(Y-\bar{Y})^3]$$
$$\mathbb{E}[(X-\bar{X})^2(Y-\bar{Y})] \quad \text{and} \quad \mathbb{E}[(X-\bar{X})(Y-\bar{Y})^2]$$

Correlation Coefficient

Correlation Coefficient

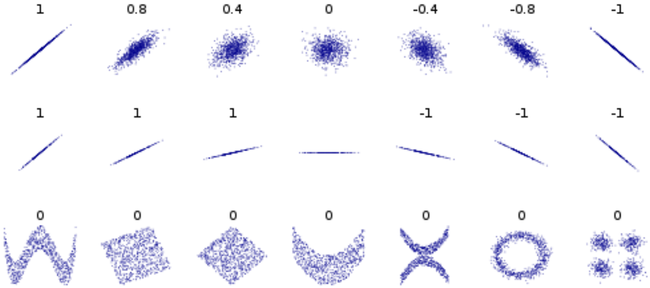
The normalized second order central moment ρ is denoted as ***correlation coefficient*** and is given by

$$\rho = \frac{\mathbb{E}[(X-\bar{X})(Y-\bar{Y})]}{\sigma_X \sigma_Y} = \frac{\mathbb{E}[(X-\bar{X})(Y-\bar{Y})]}{\sqrt{\mathbb{E}[(X-\bar{X})^2]}\sqrt{\mathbb{E}[(Y-\bar{Y})^2]}}$$

- If $\rho = 0$, then X and Y are ***uncorrelated***
- The value of ρ varies in the range $-1 \leq \rho \leq 1$

Correlation Coefficient

- The correlation coefficient varies in the range $-1 \leq \rho \leq 1$
- The correlation coefficient for different data sets

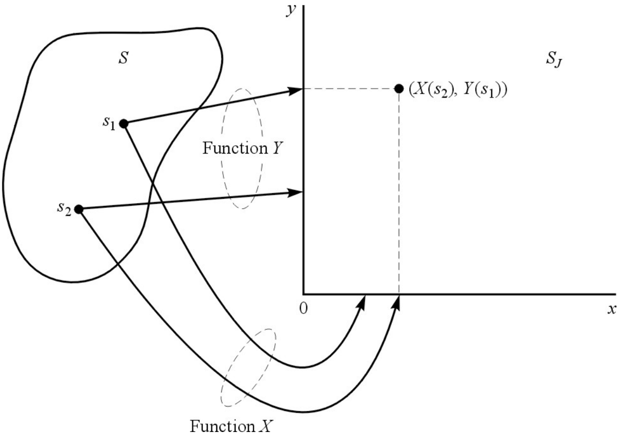


Progress...

- Last section
  - Bivariate discrete random variable
- Current section
  - Bivariate continuous random variable

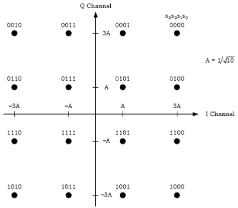
Vector Random Variable

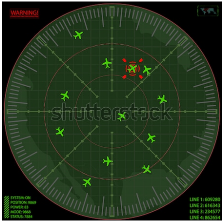
- Consider two random variables X and Y that are defined on an outcome space S

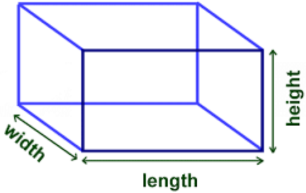


Example

- Digitally modulated signals
- Two-dimensional location on a radar screen
- Dimensions of objects over a production belt

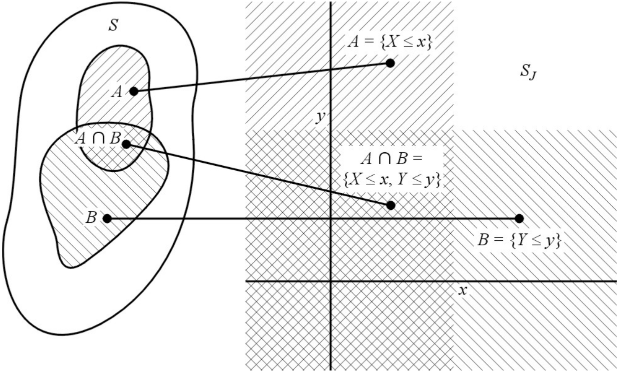






Bivariate Random Variable

- Define the events $\mathcal{A} = \{X \leq x\}$ and $\mathcal{B} = \{Y \leq y\}$



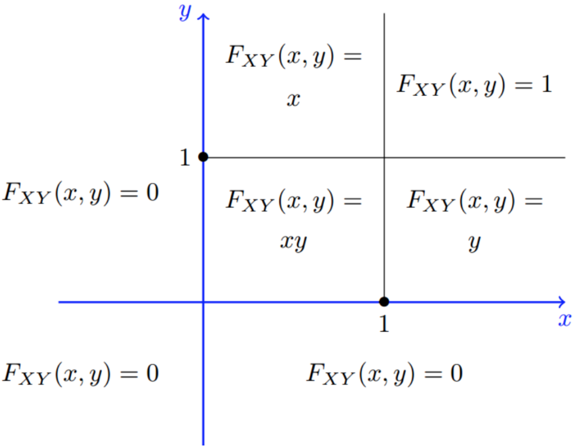
Example

- Let $X \sim Uniform(1)$ and $Y \sim Uniform(1)$ be independent random variables.
- Find the joint PDF and CDF of X and Y
- The range $R_{XY} = [0,1] \times [0,1]$

$$f_{XY}(x,y) = \begin{cases} 1; & \text{for } 0 \leq x, y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$F_{XY}(x,y) = \begin{cases} 0 & \text{for } x \leq 0 \text{ or } y \leq 0 \\ xy & \text{for } 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 1 \\ x & \text{for } 0 \leq x \leq 1 \text{ and } y \geq 1 \\ y & \text{for } x \geq 1 \text{ and } 0 \leq y \leq 1 \\ 1 & \text{for } x \geq 1 \text{ and } y \geq 0 \end{cases}$$

Example



Joint Distribution Functions

- For the bivariate case, the joint distribution function is defined as

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

Properties of Distribution Functions

- The CDF has the following properties
  1. $F_{X,Y}(-\infty,-\infty) = 0$, $F_{X,Y}(x,-\infty) = 0$ $F_{X,Y}(-\infty,y) = 0$
  2. $F_{X,Y}(\infty,\infty) = 1$
  3. $0 \leq F_{X,Y}(x,y) \leq 1$
  4. $F_{X,Y}(x,y)$ is a nondecreasing function is both x and y
  5. $\mathbb{P}(x_1 < X \leq x_2, y_1 < Y \leq y_2) =$
     $F_{X,Y}(x_2,y_2) - F_{X,Y}(x_1,y_2) - F_{X,Y}(x_2,y_1) + F_{X,Y}(x_1,y_1)$
  6. $F_{X,Y}(x,\infty) = F_X(x)$ & $F_{X,Y}(\infty,y) = F_Y(y)$

Joint Density Functions

- For the bivariate case, the joint densify function is defined as

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$$

Properties of Density Functions

- The PDF has the following properties

1. $f_{X,Y}(x,y) > 0$
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = 1$
3. $F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(w,v)dvdw$
4. $F_X(x) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f_{X,Y}(w,v)dvdw$
   $F_Y(y) = \int_{-\infty}^{y}\int_{-\infty}^{-\infty} f_{X,Y}(w,v)dwdv$
5. $\mathbb{P}(x_1 < X \leq x_2, y_1 < Y \leq y_2) = \int_{x_1}^{x_2}\int_{y_1}^{y_2} f_{X,Y}(x,y)dydx$
6. $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$
   $f_Y(y) = \int_{-\infty}^{-\infty} f_{X,Y}(x,y)dx$

- Caution: Sometimes the integration boundaries are dependent

Example

- Consider the random variables X and Y with the following joint PDF

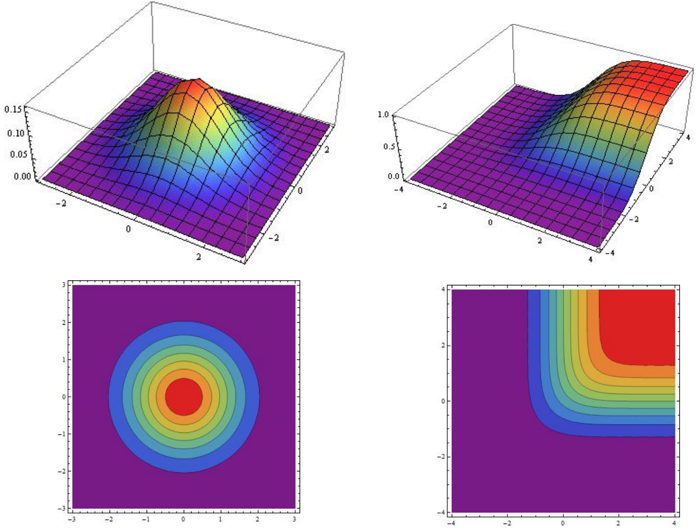
$$f_{X,Y}(x,y) = \begin{cases} be^{-x}\cos(y) & 0 \leq x \leq 2 \text{ \& } 0 \leq y \leq \frac{\pi}{2} \\ 0 & \text{elsewhere} \end{cases}$$

- Find b such that $f_{X,Y}(x,y)$ is a legitimate density function
- Solution

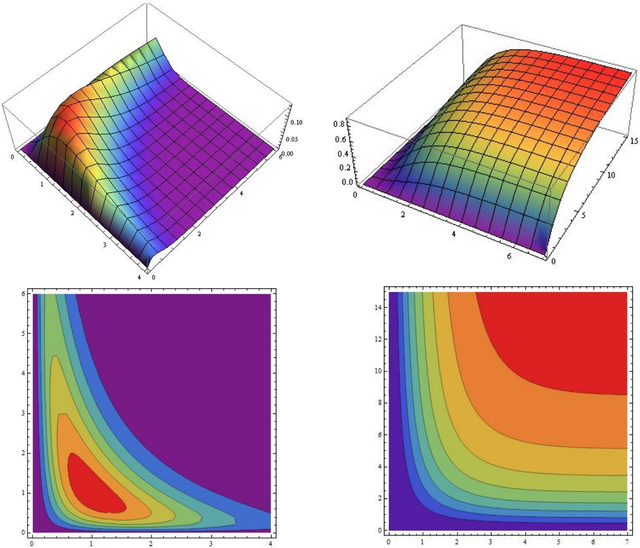
$$\begin{aligned} 1 &= \int_0^{\frac{\pi}{2}} \int_0^2 be^{-x}\cos(y)dxdy \\ &= b\int_0^2 e^{-x}dx \int_0^{\frac{\pi}{2}} \cos(y)dy \\ &= b(1-e^{-2}) \end{aligned}$$

- Hence, $b = \frac{1}{1-e^{-2}}$

Bivariate distribution



Bivariate distribution



Example

- Consider the random variables X and Y with the following joint densify function

$$f_{X,Y}(x,y) = xe^{-x(y+1)}, \quad 0 \le x, y \le \infty$$

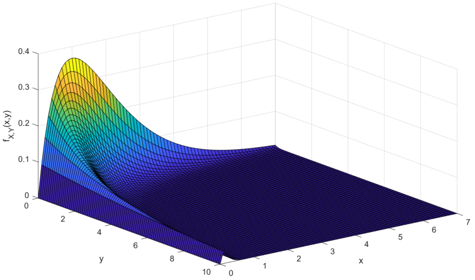
- Find the marginal PDFs of X and Y
- The marginal PDF of X is

$$\begin{aligned} f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy &= \int_0^{\infty} xe^{-x(y+1)}dy \\ &= xe^{-x}\int_0^{\infty} e^{-xy}dy = xe^{-x}\left[-\frac{e^{-xy}}{x}\right]_0^{\infty} \\ &= e^{-x}, \quad 0 \le x \le \infty \end{aligned}$$

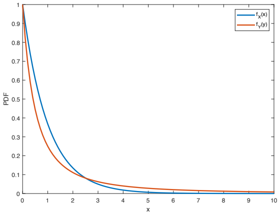
- The marginal PDF of Y is

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx = \int_0^{\infty} xe^{-x(y+1)}dx \\ &= \int_0^{\infty} xe^{-x(y+1)}dx = \left[e^{-x(y+1)}\left(-\frac{x}{(y+1)} - \frac{1}{(y+1)^2}\right)\right]_0^{\infty} \\ &= \frac{1}{(y+1)^2}, \quad 0 \le y \le \infty \end{aligned}$$

Example



0.4
0.3
0.2
0.1
0
$f_{X,Y}(x,y)$
0
2
4
6
8
10
y
0
1
2
3
4
5
6
7
x



$f_X(x)$
$f_Y(y)$
PDF
x
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
1
2
3
4
5
6
7
8
9
10

Conditional Density & Distribution Functions

- The conditional distribution function is defined as

$$F_X(x|B) = \mathbb{P}(X \leq x|B) = \frac{\mathbb{P}(X \leq x \cap B)}{\mathbb{P}(B)}$$

- The conditional density function is obtained as

$$f_X(x|B) = \frac{dF_X(x|B)}{dx}$$

- The event B can be defined as
  - Point conditioning
  - Interval conditioning

Point Conditioning

Conditional density functions: point conditioning

- The conditional distribution & density of X are

$$F_X(x|y) = \frac{\int_{-\infty}^{x} f_{X,Y}(x,v)dv}{f_Y(y)} \quad \text{and} \quad f_X(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

- The conditional distribution & density of Y are

$$F_Y(y|x) = \frac{\int_{-\infty}^{y} f_{X,Y}(x,v)dv}{f_X(x)} \quad \text{and} \quad f_Y(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

Example

- Consider the random variables X and Y with the following joint densify function

$$f_{X,Y}(x,y) = xe^{-x(y+1)}, \quad 0 \leq x, y \leq \infty$$

- Find the conditional PDF $f_Y(y|x)$
- The marginal PDF of X is

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_0^{\infty} xe^{-x(y+1)}dy \\ &= xe^{-x}\int_0^{\infty} e^{-xy}dy = xe^{-x}\left[-\frac{e^{-xy}}{x}\right]_0^{\infty} \\ &= e^{-x}; \quad 0 \leq x \leq \infty \end{aligned}$$

- The Conditional PDF of $F_Y(y|x)$ is

$$\begin{aligned} F_Y(y|x) &= \frac{f_{X,Y}(x,y)dx}{f_x(x)} = \frac{xe^{-x(y+1)}dx}{e^{-x}} \\ &= xe^{-xy}dy, \quad 0 \leq x, y \leq \infty \end{aligned}$$

Interval Conditioning

Conditional density functions: interval conditioning

- The event B is defined as $B = \{y_a \leq Y \leq y_b\}$
- The conditional density function is

$$F_X(x|y_a \leq Y \leq y_b) = \frac{\int_{y_a}^{y_b} \int_{-\infty}^{x} f_{X,Y}(w,v)dwdv}{\int_{y_a}^{y_b} f_Y(v)dv}$$

- For discrete random variable

$$F_X(x|y_a \leq Y \leq y_b) = \frac{F_{X,Y}(x,y_b) - F_{X,Y}(x,y_a)}{F_Y(y_b) - F_Y(y_a)}$$

$$f_X(x|y_a \leq Y \leq y_b) = \frac{\int_{y_a}^{y_b} f_{X,Y}(w,v)dwdv}{\int_{y_a}^{y_b} f_Y(v)dv}$$

Example

- Consider the random variables X and Y with the following joint densify function

$$f_{X,Y}(x,y) = xe^{-x(y+1)}, \quad {\scriptstyle 0 \leq x, y \leq \infty}$$

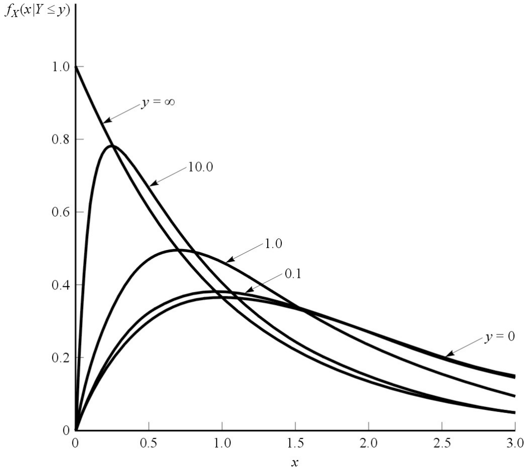
- Find the conditional PDF $f_Y(x|Y < y)$
- The marginal CDF of Y is

$$\begin{aligned} F_Y(y) &= \int_{-\infty}^{y}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = \int_{0}^{y}\int_{0}^{\infty} xe^{-x(y+1)}dxdy \\ &= \int_{0}^{y}\int_{0}^{\infty} xe^{-x(y+1)}dxdy = \int_{0}^{y}\left[e^{-x(y+1)}\left(-\frac{x}{(y+1)} - \frac{1}{(y+1)^2}\right)\right]_0^{\infty} dy \\ &= \int_{0}^{y} \frac{1}{(y+1)^2}dy = \frac{y}{(y+1)}, \quad 0 \leq y \leq \infty \end{aligned}$$

$$\begin{aligned} f_X(x|Y \leq y) &= \frac{\int_{-\infty}^{y} f_{X,Y}(x,v)dv}{F_Y(y)} = \frac{(y+1)u(x)xe^{-x}}{y}\int_0^y e^{-xv}dv \\ &= \frac{(y+1)e^{-x}\left(1 - e^{-xy}\right)}{y}, \quad 0 \leq x, y \leq \infty \end{aligned}$$

Example

- The conditional density $f_X(x|Y < y)$



Statistical Independence

- Two events A and B are statistically independent if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

- The statistical independence for random variables can be stablished be defining $A = \{X \leq x\}$ and $B = \{Y \leq y\}$

Statistical Independence

- Two random variables X and Y are statistically independent iff

$$F_{X,Y}(x, y) = F_X(x)F_Y(y)$$

- Equivalently,

$$f_{X,Y}(x, y) = f_X(x)f_Y(y)$$

- Hence,

$$f_X(x|y) = f_X(x) \quad \text{and} \quad f_Y(y|x) = f_Y(y)$$

Conditional Expectation

- Conditional expectation is similar to the ordinary single random variable expectation but using the conditional PDF

Conditional Expectation

- Two X and Y be two discrete random variable then,

$$\mathbb{E}[X|Y=y] = \int_x x f_{X|Y}(x|y) dx$$

Law of Total Probability

Law of Total Probability

- Since the outcomes of X and Y are disjoint, we have

$$f_X(x) = \int_y f_{X,Y}(x,y)dy = \int_y f_{X|Y}(x|y)f_Y(y)dy$$

- Applying the same concept to expectation

$$\mathbb{E}[X] = \int_y \mathbb{E}[X|Y=y]f_Y(y)dy = \mathbb{E}[\mathbb{E}[X|Y]]$$

- This is denoted as the law of iterated expectation

Example

- Consider the random variables X and Y with the following joint densify function

$$f_{X,Y}(x,y) = xe^{-x(y+1)}, \quad 0 \leq x, y \leq \infty$$

- The marginal PDF of X is

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = u(x)e^{-x}$$

- The marginal PDF of Y is

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx = \frac{u(y)}{(y+1)^2}$$

- X and Y are not independent

$$f_y(x)f_Y(y) = \frac{e^{-x}}{(y+1)^2} \neq f_{X,Y}(x,y)$$

Example

- Consider the random variables X and Y with the following joint densify function

$$f_{X,Y}(x,y) = \frac{e^{-\frac{x}{4}-\frac{y}{3}}}{12}, \quad 0 \leq x, y \leq \infty$$

- The marginal PDF of X is

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \frac{e^{-\frac{x}{4}}}{4}, \quad 0 \leq x \leq \infty$$

- The marginal PDF of Y is

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx = \frac{e^{-\frac{y}{3}}}{3}, \quad 0 \leq y \leq \infty$$

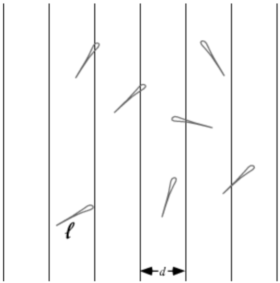
- X and Y are independent

$$f_y(x)f_Y(y) = \frac{e^{-\frac{x}{4}-\frac{y}{3}}}{12} = f_{X,Y}(x,y)$$

Example: Buffon’s needle

Buffon’s needle

- Consider a 2 dimensional space with parallel lines with interspaces of d
- A needle of length ($l < d$) is thrown at random over the space
- Find the probability that the needle intersects one of the lines



Solution: Buffon's needle

- Let X denote the distance from the needle's center to the nearest line
- Let θ be the acute angle between the needle and a parallel line passing through the needle's center
- Then

$$f_X(x) = \frac{2}{d}; \quad 0 \leq x \leq \frac{d}{2} \quad \text{and} \quad f_\Theta(\theta) = \frac{2}{\pi}; \quad 0 \leq \theta \leq \frac{\pi}{2}$$

$$f_{X,\Theta}(x,\theta) = f_X(x) \times f_\Theta(\theta) = \frac{4}{\pi d}; \quad \begin{array}{l} 0 \leq x \leq \frac{d}{2} \\ 0 \leq \theta \leq \frac{\pi}{2} \end{array}$$

- The probability of line intersection is

$$\mathbb{P}(X < \frac{l}{2}\sin(\Theta)) = \int_0^{\frac{\pi}{2}} \int_0^{\frac{l}{2}\sin(\Theta)} \frac{4}{\pi d} dx d\theta = \int_0^{\frac{\pi}{2}} \frac{2l\sin(\Theta)}{\pi d} d\theta$$
$$= \frac{2l}{\pi d}$$

Example

Break a sick twice

- Break a stick of length ℓ twice
- First break at X: uniform in $[0, \ell]$
- Second break at Y: uniform from $[0, X]$
- Find the marginal distribution of Y

- The joint PDF of X and Y is

$$\begin{aligned} f_{X,Y}(x,y) &= f_X(x) f_{Y|X}(y|x) \\ &= \frac{1}{x\ell}; \qquad 0 \leq y \leq x \leq \ell \end{aligned}$$

- The marginal PDF of Y is

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx = \int_{y}^{\ell} \frac{1}{x\ell} dx \\ &= \frac{\ln(\ell) - \ln(y)}{\ell} = \frac{1}{\ell} \ln\left(\frac{\ell}{y}\right) \quad 0 \leq y \leq \ell \end{aligned}$$

Expected value

Expected value of a function of random variables

- Let $g(X, Y)$ be a function that involve two continuous random variables X and Y. Then,

$$\bar{g} = \mathbb{E}[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) dx dy$$

- For $g(X, Y) = X^n Y^k$, we have

$$\mathbb{E}[X^n Y^k] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^n y^k f_{X,Y}(x, y) dx dy$$

- For $g(X, Y) = (X - \bar{X})^n (Y - \bar{Y})^k$. we have

$$\mathbb{E}[(X - \bar{X})^n (Y - \bar{Y})^k] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{X})^n (y - \bar{Y})^k f_{X,Y}(x, y) dx dy$$

Joint Moments

- First order moments about the origin

$$m_{1,0} = \mathbb{E}[X] \quad \text{and} \quad m_{0,1} = \mathbb{E}[Y]$$

- Second order moments about the origin

$$m_{2,0} = \mathbb{E}[X^2] \quad \text{and} \quad m_{0,2} = \mathbb{E}[Y^2] \quad \text{and} \quad m_{1,1} = \mathbb{E}[XY]$$

- Third order moments about the origin

$$m_{3,0} = \mathbb{E}[X^3] \quad \text{and} \quad m_{0,3} = \mathbb{E}[Y^3]$$
$$m_{2,1} = \mathbb{E}[X^2 Y] \quad \text{and} \quad m_{1,2} = \mathbb{E}[XY^2]$$

Correlation

Correlation

The second moment about the origin $m_{1,1} = R_{X,Y}$ is of special interest and is denoted as the correlation of X and Y

$$\mathbb{E}[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{X,Y}(x,y)dxdy$$

- If $R_{XY} = \mathbb{E}[X]\mathbb{E}[Y]$, then X and Y are ***uncorrelated***
- Uncorrelation does not imply independence
- Independence implies uncorrelation
- If $R_{XY} = 0$, then X and Y are ***orthogonal***

Proof: Independence implies uncorrelation

- Let X and Y be two independent random variables. Then, of X and Y

$$\begin{aligned}
\mathbb{E}[XY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{X,Y}(x,y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)dxdy \\
&= \left(\int_{-\infty}^{\infty} xf_X(x)dx\right)\left(\int_{-\infty}^{\infty} yf_Y(y)dy\right) \\
&= \mathbb{E}[X]\mathbb{E}[Y]
\end{aligned}$$

Covariance

Covariance

The second order central moment $C_{X,Y}$ is of special interest and is denoted as the covariance of X and Y, which is given by

$$\mathbb{E}[(X-\bar{X})(Y-\bar{Y})] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\bar{X})(y-\bar{Y})f_{X,Y}(x,y)dxdy$$
$$= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

- If $C_{XY} = 0$, then X and Y are ***uncorrelated***
- If $C_{XY} = -\mathbb{E}[X]\mathbb{E}[Y]$, then X and Y are ***orthogonal***

Correlation Coefficient

Correlation Coefficient

The normalized second order central moment $\rho = \frac{\mu_{11}}{\sigma_X \sigma_Y}$ is denoted as ***correlation coefficient*** and is given by

$$\begin{aligned} \rho &= \frac{\mathbb{E}[(X - \bar{X})(Y - \bar{Y})]}{\sigma_X \sigma_Y} \\ &= \frac{\mathbb{E}[(X - \bar{X})(Y - \bar{Y})]}{\sqrt{\mathbb{E}[(X - \bar{X})^2]}\sqrt{\mathbb{E}[(Y - \bar{Y})^2]}} \end{aligned}$$

- If $\rho = 0$, then X and Y are ***uncorrelated***
- The value of ρ varies in the range $-1 \leq \rho \leq 1$

Table of contents

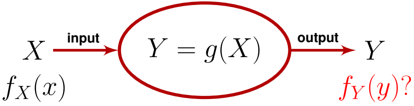
Progress

- Last section
  - Bivariate discrete random variables
  - Bivariate continuous random variables
- Current section
  - Transformation of multiple random variables
    - Single function
    - Multiple functions
  - Linear transformation of gaussian random vectors

Revision

- Transformation of a single random variable

$$Y = g(X)$$



Theorem

- For monotonic functions, let $x_1 = g^{-1}(x)$, then the density function of Y is given by

$$f_Y(y) = \frac{f_X(x_1)}{|g'(x_1)|}$$
$$= f_X\left(g_i^{-1}(y)\right)\left|\frac{dg_i^{-1}(y)}{dy}\right|$$

Functions of two RVs

Theorem

- Let

$$\begin{array}{l} W = g_1(X,Y) \\ V = g_2(X,Y) \end{array} \implies \begin{array}{l} X = g_1^{-1}(W,Y) \\ Y = g_2^{-1}(W,V) \end{array}$$

- The joint density function of W, V is given by

$$f_{W,V}(w,v) = f_{X,Y}\left(g_1^{-1}(w,v), g_2^{-1}(w,v)\right)|J|$$

where $|J|$ is the absolute value of the determinant of the Jacobian matrix

Jacobian matrix

Definition: Jacobian matrix

- Consider the functions

$$\begin{gathered} g_1(x,y) \\ g_2(x,y) \end{gathered}$$

- The Jacobian matrix for these two functions is given by

$$J = \begin{bmatrix} \frac{\partial g_1(x,y)}{\partial x} & \frac{\partial g_1(x,y)}{\partial y} \\ \frac{\partial g_2(x,y)}{\partial x} & \frac{\partial g_2(x,y)}{\partial y} \end{bmatrix}$$

- Then $|J|$ is the absolute value to the determinant of the Jacobean matrix

Example 3

- Let $Y = \frac{X_1}{X_2}$, where X_1 and X_2 are two positive random variables
- Find the density function of Y using the Jacobian method
- Define the auxiliary $Z = X_2$
- Then, $X_1 = g_1^{-1}(Y, Z) = YZ$ and $X_2 = g_2^{-1}(Y, Z) = Z$
- The Jacobian is given by

$$J = \begin{bmatrix} \frac{\partial g_1^{-1}(Y,Z)}{\partial Y} & \frac{\partial g_1^{-1}(Y,Z)}{\partial Z} \\ \frac{\partial g_2^{-1}(Y,Z)}{\partial Y} & \frac{\partial g_2^{-1}(Y,Z)}{\partial Z} \end{bmatrix} = \begin{bmatrix} Z & Y \\ 0 & 1 \end{bmatrix} \implies |J| = Z$$

- The joint density function of Y and Z is given by

$$f_{Y,Z}(z, y) = f_{X_1,X_2}(g_1^{-1}(y, z), g_2^{-1}(y, z))|J| = z f_{X_1,X_2}(yz, z)$$

- Then, the marginal density function of Y is given by

$$f_Y(z) = \int_0^{\infty} z f_{X_1,X_2}(yz, z) dz$$

Example 4

- Let $Y_1 = aX_1 + bX_2$ and $Y_2 = cX_1 + dX_2$
- Find the joint density function of Y_1 and Y_2
- Simultaneously solving the two linear equations for X_1 and X_2, we have

$$X_1 = g_1^{-1}(\mathbf{y}) = \frac{dY_1 - bY_2}{ad - bc} \quad \text{and} \quad X_2 = g_2^{-1}(\mathbf{y}) = \frac{-cY_1 + aY_2}{ad - bc}$$

- The Jacobian is given by

$$J = \begin{bmatrix} \frac{\partial g_1^{-1}(\mathbf{y})}{\partial Y_1} & \frac{\partial g_1^{-1}(\mathbf{y})}{\partial Y_2} \\ \frac{\partial g_2^{-1}(\mathbf{y})}{\partial Y_1} & \frac{\partial g_2^{-1}(\mathbf{y})}{\partial Y_2} \end{bmatrix} = \begin{bmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix} \implies |J| = \frac{1}{|ad - bc|}$$

- The joint density function of Y and Z is given by

$$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) &= f_{X_1,X_2}\left(g_1^{-1}(\mathbf{y}), g_2^{-1}(\mathbf{y})\right) |J| \\ &= \frac{1}{|ad - bc|} f_{X_1,X_2}\left(\frac{dY_1 - bY_2}{ad - bc}, \frac{-cY_1 + aY_2}{ad - bc}\right) \end{aligned}$$

Questions?