LECTURE NOTE

ON

PROBABILITY AND STATISTICS 1

BY

ASSIST. PRF. DR. MUSTAFA I. NAIF

DEPARTMENT OF MATHEMATICS
COLLEGE OF EDUCATION FOR PURE SCIENCE
UNIVERISTY OF ANBAR

LECTURE 1#

- Contents :-

A:-

-The Basic principle of counting .

-Permutations .

-Combinabions .

LECTURE 1#

- Contents :-

B:-

-Axioms of probabilly .

-some simple propositions .

-sample space and events .

LECTURE 1#

- Contents :-

C:-

-Random variables.

-Discrete Random variables .

-Expected value .

-Variance .

-Probability distribution function

- Discrete distributions

LECTURE 1#

- Contents :-

D:-Continuous random variables

-Expectation and variance of Continuous random variables
-Continuous distribution .

-The distribution functions.

E:-conditional probability and independence

F:-Jointly distribution

G:-Properties of Expectation

References

- Mathematical Statistics with Applications. D. D. Wackerly, William Mendenhall and Richard L. Scheaffer, seven edition, 2008

- Probability and Statistics. Morris H. DeGroot and Mark J. Schervish, Fourth Edition,2012

- A FIRST COURSE IN PROBABILITY. Sheldon Ross, Ninth Edition, 2014

The Basic principle of counting

The Fundamental Counting Principle is a way of determining the number of possible ways that we can perform two or more operations together. If operations were being performed independent of each other instead of together, then we would NOT use the Fundamental Counting Principle.
The Basic or Fundamental Counting Principle can be used to find the number of possibilities when given several groups. **How?** By Multiply the number of elements in each group together.

Def: Multiplication principle (Fundamental Principle of Counting):
Suppose an event E can occur in m different ways and associated with each way of occurring of E, another event F can occur in n different ways, then the total number of occurrence of the two events in the given order is $m \times n$.

Def: Addition principle
If an event E can occur in m ways and another event F can occur in n ways, and suppose that both can not occur together, then E or F can occur in $m + n$ ways.

Example: Ice cream comes in either a cup or a cone and the flavors available are chocolate, strawberry and vanilla.

```
                          / chocolate cup
         / chocolate <
        /                 \ chocolate cone
       /
      /                  / strawberry cup
<-- strawberry <
     \                   \ strawberry cone
      \
       \                / vanilla cup
        \ vanilla <
                        \ vanilla cone
```

this diagram is called a tree diagram and shows all of the possibilities. The tree diagram could also be arranged in another way. Both diagrams have 6 total outcomes.

```
            / chocolate cup
           /
     / cup <-- strawberry cup
    /      \
   /        \ vanilla cup
  /
<
 \
  \          / chocolate cone
   \        /
    \ cone <-- strawberry cone
            \
             \ vanilla cone
```

Note: If there are more than two outcomes, continue to multiply the possibilities together to determine the total outcomes.

To determine the total number of outcomes, multiply the number of possibilities of the first characteristic times the number of possibilities of the second characteristic. **In the example above, multiply 3 times 2 to get 6 possible outcomes.**

Example: **How many different license plates are there altogether?**

Look at what's used to make a plate:

LETTER	LETTER	LETTER	NUMBER	NUMBER	NUMBER

For each of the letters we have 26 choices. For each of the numbers we have 10 choices.

The number of ways to pick the first letter	The number of ways to pick the second letter	The number of ways to pick the third letter	The number of ways to pick the first number	The number of ways to pick the second number	The number of ways to pick the third number
26	26	26	10	10	10

The Fundamental Counting Principle says that:

The total number of ways to fill the six spaces on a license plate is

26** x **26** x **26** x **10** x **10** x **10

which equals 17,576,000

Example 3: **In a class, there are 27 boys and 14 girls. The teacher wants to select 1 boy and 1 girl to represent the class for a function. In how many ways can the teacher make this selection?**

Solution:

The teacher is to perform two operations:

1)Selecting a boy from among the 27 boys and

2) Selecting a girl from among 14 girls.

The first of these can be done in 27 ways and second can be performed in 14 ways. So, By the fundamental principle of counting, the required number of ways is:

$\mathbf{27 \times 14 = 378}$

Example 4:

1) How many numbers are there between 99 and 1000 having 7 in the units place?

2) How many numbers are there between 99 and 1000 having at least one of their digits 7?

Solution:

1) First note that all these numbers have three digits. 7 is in the unit's place. The middle digit can be any one of the 10 digits from 0 to 9. The digit in hundred's place can be any one of the 9 digits from 1 to 9. Therefore, by the fundamental principle of counting, there are 10 × 9 = 90 numbers between 99 and 1000 having 7 in the unit's place.

1-9	0-9	7
hundred's place	**The middle's place**	**the unit's place**

2) Total number of 3 digit numbers having at least one of their digits as 7 = (Total numbers of three digit numbers) – (Total number of 3 digit numbers in which 7 does not appear at all). = (9 × 10 × 10) – (8 × 9 × 9) = 900 – 648 = 252.

1-9	0-9	0-9	-	1-9	0-9	0-9
9	**10**	**10**		**8**	**9**	**9**

NOTE:

If you have a problem where you can repeat objects, then you must use the Fundamental Counting Principle; you cannot us Permutations or Combinations.

See you next Lecture

LECTURE 2#

Ex.1: How many numbers consist of two digits can be create from {1,2,3,4,5} if:

1) We can repeat the number (with repetition).

2) We can not repeat the number(without repetition)

Sol:

1) The number of ways are $5 \times 5=25$.
2) The number of ways are $5 \times 4=20$.

Ex.2: How many ways that possible to set 5 persons on 5 chairs?

Sol: $5! = 5 \times 4 \times 3 \times 2 \times 1=120$.

Permutations

- A permutation is an arrangement of objects in a definite order.

1)Permutation of n different objects: (without rep.)

-The number of permutations of n objects taken all at a time, denoted by the symbol P_n is given by:

$$P_n = n!$$

-The number of permutations of n objects taken r at a time, where 0 < r ≤ n, denoted by P_r^n is given by:

$$P_r^n = \frac{n!}{(n-r)!}$$

$0! = 1$

Permutations

EX.3 :As we saw in EX.1 part 2 and this is for case without repetition.

EX.4 : The number of possible 3 per. of 5 objects is:

$$P_3^5 = \frac{5!}{(5-3)!} = 60$$

2) When repetition of objects is allowed:

-The number of permutations of n things taken all at a time, when repetition of objects is allowed is n^n

- The number of permutations of n objects, taken r at a time, when repetition of objects is allowed, is n^r

Permutations

EX.5 : box containing 8 different balls. You have drawn 3 balls one after the other. Find the number of ways to draw the three balls if the drawing is:

1- without repetition

2 - With repetition

Sol: 1) Number of ways to draw the first ball= 8

Number of ways to draw the second ball =7

Number of ways to draw the third ball =6

So, the total number of ways to draw the three balls = 8 X 7 X6=336

2) Number of ways to draw the first ball=8

Number of ways to draw the second ball=8

Number of ways to draw the third ball=8

the total number = 8X8X8=512

Permutations

EX.6 : If we have the numbers 1, 2,3,4,5,6,7 .

1- How many 3-digit number can be formed from the previous numbers without repeating?

2- How many numbers consisting of 3 odd numbers can be formed from the previous numbers without repeating?

3- How many 3-digit number starting with 4 can be formed from the previous numbers without repeating?

4- How many 3-digit number greater than 300 can be formed from the previous numbers without repeating?

Permutations

Sol :

1- $P_r^n = \frac{n!}{(n-r)!}$ = 7 X 6X 5= 210 ,Note: $P_r^n = \frac{n!}{(n-r)!}$=n(n-1)…(n-r+1)

2- No. of ways= 4 X 6 X 5 =120

3- No. of ways = 1 X 6 X 5=30 (نفس المثال في المحاضرة السابقة)

4- No. of ways = 5 X 6 X 5= 150 (نفس المثال في المحاضرة السابقة)

Permutations

Homework:

1- If $P_r^{12} = 1320$, find r

2-If $P_r^{8} = 6720$, find (r+1)!

3- If $P_4^{n} = 14\ X\ \ P_3^{(2-n)}$, find n

4- In how many ways can 5 children be arranged in a line such that

(1) two particular children of them are always together

(2) two particular children of them are never together.

5-How many ways can five male and five female students sit in a row of ten seats so that the students are side by side and the female students side by side

See you next Lecture

LECTURE 3#

Combinations

A combination is a selection of some or all of a number of different objects where the order of selection is immaterial. The number of selections of r objects from the given n objects is denoted by C_r^n , and is given by:

$$C_r^n = \frac{n!}{r!\,(n-r)!} = \binom{n}{r}$$

Notes:

1-Use permutations if a problem calls for the number of arrangements of objects and different orders are to be counted.

2- Use combinations if a problem calls for the number of ways of selecting objects and the order of selection is not to be counted.

Combinations

Ex.1: In how many ways a committee consisting of 2 persons, can be chosen from 5 persons ?

Sol: $C_2^5 = \frac{5!}{2!(5-2)!} = \binom{5}{2} = 10.$

Note:

1- $C_r^n = \frac{n!}{r!(n-r)!} = \frac{P_r^n}{r!}$

2- $C_r^n = C_{n-r}^n$

3- $If\ C_r^n = C_s^n \ \rightarrow$ r=s or r + s = n

4- $C_r^{n+1} = C_{r-1}^n + C_r^n$

5- n$C_{r-1}^{n-1} = (n - r + 1)C_{r-1}^n$

Combinations

EX.3 : How many different two committees are possible made from a group of 12 people such that each committee consistent of 3 persons and there is no common person between these two committees?

Sol: The possible ways for selecting the first committee is:

$$C_3^{12} = 220 \, ways$$

The possible ways for selecting the second committee is:

$$C_3^{9} = 84 \, ways$$

So, the total required ways are: 220 X 84=18480

EX.3: In how many ways a committee consisting of 3 men and 2 women, can be chosen from 7 men and 5 women? H.W (150)

Combinations

EX.4: If $C_5^n = C_3^n$, find C_2^n ?

Sol.: By Note (3) , $C_5^n = C_3^n$ $\longrightarrow$ n=5+3=8.

Therefore, $C_2^8 = 28.$

EX.5: If $C_{r+2}^{12} = C_{2r+1}^{12}$, find r ? You try to find r

Note: we denote by $C'_{n,r}$, the number of possible combinations with repetition of r objects from n, where, $C'_{n,r} = C_r^{n+r-1} = \binom{n+r-1}{r} = \frac{(n+r-1)!}{(n-1)!r!}$

Combinations

EX.6: the number of possible combination with repetition of 3 objects from 5 is:

$$C'_{5,3} = 35.$$

EX.7: How much can you distribute 12 different books to three students, so that the first student takes 5 books, the second student takes 4 books, and the third student takes 3 books

Sol: $C_5^{12} \times C_4^7 \times C_3^3$ =792 X 35 X 1= 27720

عدد طرق اختيار الطالب: الثالث الثاني الاول

Homework

1- How many ways can committees be formed from among 18 students, so that the first committee consists of 3 students, the second committee 4 students, and the third committee consists of 6 students.

2- All the letters of the word ‘EAMCOT’ are arranged in different possible ways. What is the number of such arrangements in which no two vowels are adjacent to each other ?

3- In an examination there are three multiple choice questions and each question has 4 choices. What is the number of ways in which a student can fail to get all answer correct?

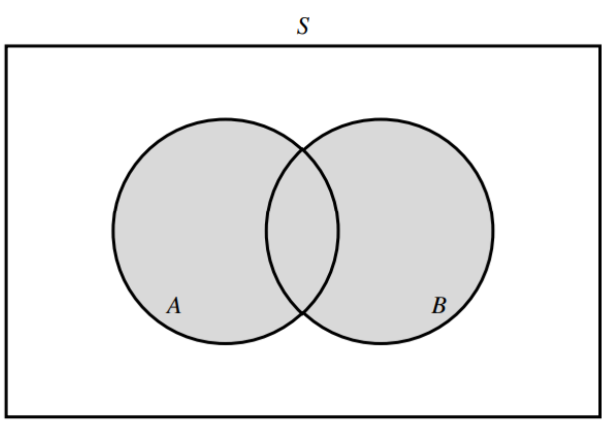
See you next Lecture

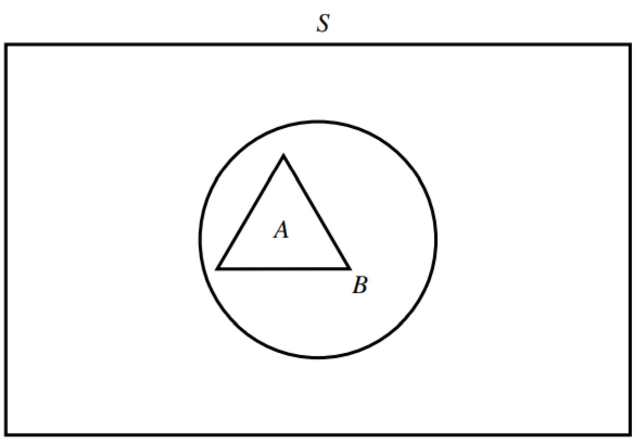
LECTURE 4#

A Review of Set Notation

- Let S denote the set of all elements under consideration; that is, S is the universal set.
- For any two sets A and B, we will say that A is a subset of B, or A is contained in B (denoted $A \subset B$), if every point in A is also in B.
- The null, or empty, set, denoted by $\emptyset$, is the set consisting of no points. Thus, $\emptyset$ is a subset of every set.
- The union of A and B, denoted by $A \cup B$, is the set of all points in A or B or both. That is, the union of A and B contains all points that are in at least one of the sets.

Venn diagram for $A \cup B$



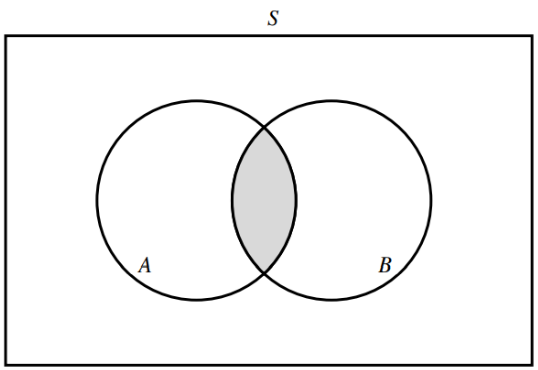


Venn diagram for $A \subset B$

A Review of Set Notation

- The intersection of A and B, denoted by $A \cap B$ or by AB, is the set of all points in both A and B.

Venn diagram for AB

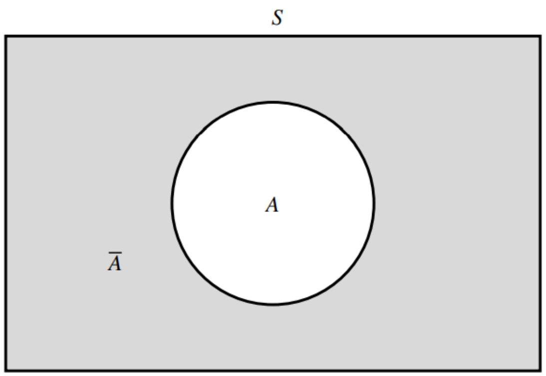


- If A is a subset of S, then the complement of A, denoted by $\bar{A}$, is the set of points that are in S but not in A.
- Two sets, A and B, are said to be disjoint, or mutually exclusive, if $A \cap B = \emptyset$.

Ex. 1: let S denote the set of all possible numerical observations for a single toss of a die. That is $S = \{1, 2, 3, 4, 5, 6\}$. Let $A = \{1, 2\}$, $B = \{1, 3\}$, and $C = \{2, 4, 6\}$.

Then $A \cup B = \{1, 2, 3\}$, $A \cap B = \{1\}$, and $\bar{A} = \{3, 4, 5, 6\}$. Also, note that B and C are mutually exclusive, whereas A and C are not.

Venn diagram for $\bar{A}$



A Review of Set Notation

- The distributive laws are given by

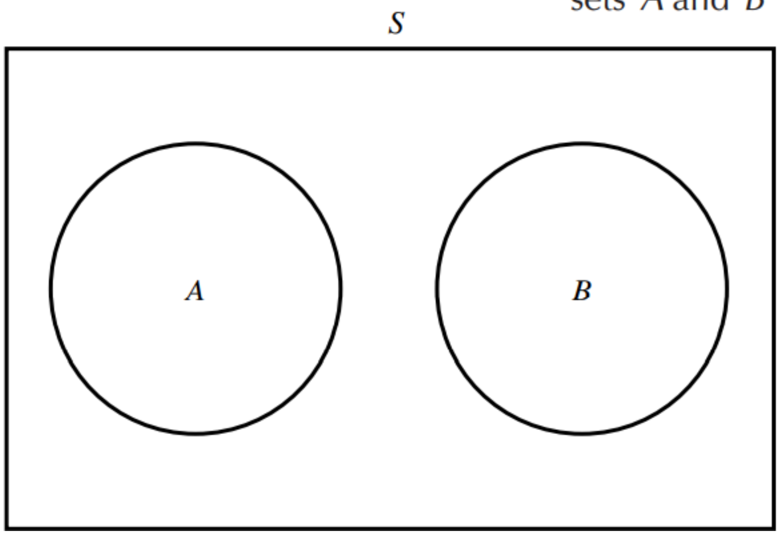
  A ∩ (B ∪ C) = (A ∩ B) ∪ (A ∩ C),

  A ∪ (B ∩ C) = (A ∪ B) ∩ (A ∪ C),

- DeMorgan's laws:

$$\overline{(A \cap B)} = \overline{A} \cup \overline{B} \quad \text{and} \quad \overline{(A \cup B)} = \overline{A} \cap \overline{B}.$$

Venn diagram for mutually exclusive sets A and B



Probability

Def. 1:An experiment is the process by which an observation is made.

Examples of experiments include coin and die tossing.

Def 2: A simple event is an event that cannot be decomposed. Each simple event corresponds to one and only one sample point. The letter E with a subscript will be used to denote a simple event or the corresponding sample point.

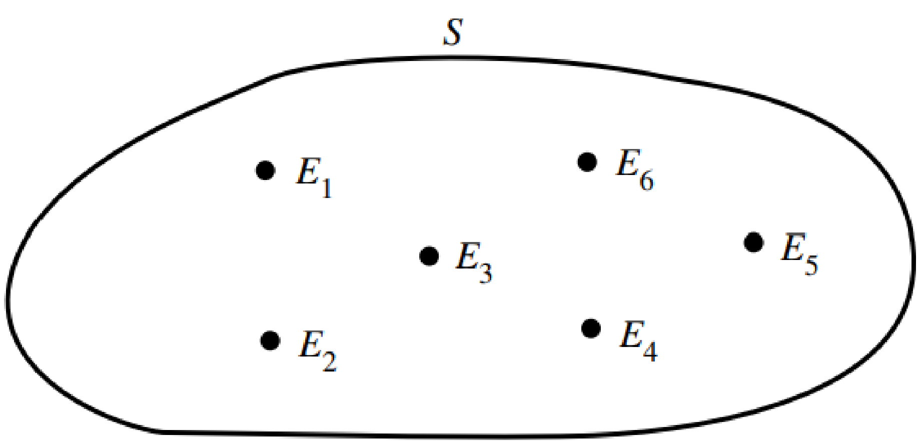
Probability

Some events associated with a single toss of a balanced die are these:

A: Observe an odd number.
B: Observe a number less than 5.
C: Observe a 2 or a 3.
E_1: Observe a 1.
E_2: Observe a 2.
E_3: Observe a 3.
E_4: Observe a 4.
E_5: Observe a 5.
E_6: Observe a 6.

- Event A, which can be decomposed into three other events, is called a compound event.
- the events E1, E2, E3, E4, E5, and E6 cannot be decomposed and are called simple events.
- An event is simple if it consists of just a single outcome, and is compound otherwise
- Because sets are collections of points, we associate a distinct point, called a sample point

Venn diagram for the sample space associated with the die-tossing experiment



Probability

- The sample space associated with an experiment is the set consisting of all possible sample points (outcomes). A sample space will be denoted by S.

EX.1: If I toss a coin three times and record the result, the sample space is

S = {HHH,HHT,HTH,HTT,THH,THT,T TH,TTT}

- An event is a subset of S. We can specify an event by listing all the outcomes that make it up.

EX. 2: In the above example, let A be the event 'more heads than tails' and B the event 'heads on last throw'. Then

A = {HHH,HHT,HTH,THH}, B = {HHH,HTH,THH,TTH}.

So, if all outcomes are equally likely, we have P(A) = $\frac{No.A}{No.S}$.

In our example, both A and B have probability P(A) =P(B)= $\frac{4}{8} = 0.5$.

Probability

- In Ex. 2, A and B are compound events, while the event 'heads on every throw' is simple (as a set, it is {HHH}).
- If A = {a} is a simple event, then the probability of A is just the probability of the outcome a, and we usually write P(a), which is simpler to write than P({a}).
- Note that a is an outcome, while {a} is an event, indeed a simple event. We can build new events from old ones:

- $A \cup B$ (read 'A union B') consists of all the outcomes in A or in B (or both!)
- $A \cap B$ (read 'A intersection B') consists of all the outcomes in both A and B;
- $A \setminus B$ (read 'A minus B') consists of all the outcomes in A but not in B;
- A' (read 'A complement') consists of all outcomes not in A (that is, $\mathcal{S} \setminus A$);
- $\emptyset$ (read 'empty set') for the event which doesn't contain any outcomes.

Probability

Remember that an event is a subset of the sample space S. A number of events, say $A_1, A_2, \ldots$, are called *mutually disjoint* or *pairwise disjoint* if $A_i \cap A_j = \emptyset$ for any two of the events A_i and A_j; that is, no two of the events overlap.

According to Kolmogorov's axioms, each event A has a probability $P(A)$, which is a number. These numbers satisfy three axioms:

According to Kolmogorov's axioms, each event A has a probability $P(A)$, which is a number. These numbers satisfy three axioms:

Axiom 1: For any event A, we have $P(A) \geq 0$.

Axiom 2: $P(S) = 1$.

Axiom 3: If the events $A_1, A_2, \ldots$ are pairwise disjoint, then

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$$

Next Lecture: You can prove simple properties of probability from the axioms.

See you next Lecture

LECTURE 5#

Proving things from the axioms

Proposition 1.1 *If the event A contains only a finite number of outcomes, say* $A = \{a_1, a_2, \ldots, a_n\}$, *then*

$$P(A) = P(a_1) + P(a_2) + \cdots + P(a_n).$$

To prove the proposition, we define a new event A_i containing only the outcome a_i, that is, $A_i = \{a_i\}$, for $i = 1, \ldots, n$. Then $A_1, \ldots, A_n$ are mutually disjoint

(each contains only one element which is in none of the others), and $A_1 \cup A_2 \cup \cdots \cup A_n = A$; so by Axiom 3a, we have

$$P(A) = P(a_1) + P(a_2) + \cdots + P(a_n).$$

Corollary 1.2 *If the sample space* $\mathcal{S}$ *is finite, say* $\mathcal{S} = \{a_1, \ldots, a_n\}$, *then*

$$P(a_1) + P(a_2) + \cdots + P(a_n) = 1.$$

For $P(a_1) + P(a_2) + \cdots + P(a_n) = P(\mathcal{S})$ by Proposition 1.1, and $P(\mathcal{S}) = 1$ by Axiom 2.

Note: Notice that once we have proved something, we can use it on the same basis as an axiom to prove further facts.

Proposition 1.3 $P(A') = 1 - P(A)$ *for any event* A.

Let $A_1 = A$ and $A_2 = A'$ (the complement of A). Then $A_1 \cap A_2 = \emptyset$ (that is, the events A_1 and A_2 are disjoint), and $A_1 \cup A_2 = \mathcal{S}$. So

$$\begin{aligned} P(A_1) + P(A_2) &= P(A_1 \cup A_2) && \text{(Axiom 3)} \\ &= P(\mathcal{S}) \\ &= 1 && \text{(Axiom 2).} \end{aligned}$$

So $P(A) = P(A_1) = 1 - P(A_2)$.

Corollary 1.4 $P(A) \leq 1$ *for any event* A.

For $1 - P(A) = P(A')$ by Proposition 1.3, and $P(A') \geq 0$ by Axiom 1; so $1 - P(A) \geq 0$, from which we get $P(A) \leq 1$.

Remember that if you ever calculate a probability to be less than 0 or more than 1, you have made a mistake!

Corollary 1.5 $P(\emptyset) = 0$.

For $\emptyset = \mathcal{S}'$, so $P(\emptyset) = 1 - P(\mathcal{S})$ by Proposition 1.3; and $P(\mathcal{S}) = 1$ by Axiom 2, so $P(\emptyset) = 0$.

Proposition 1.6 *If $A \subseteq B$, then $P(A) \leq P(B)$.*

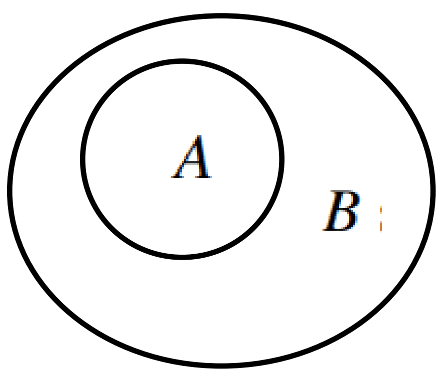
This time, take $A_1 = A$, $A_2 = B \setminus A$. Again we have $A_1 \cap A_2 = \emptyset$ (since the elements of $B \setminus A$ are, by definition, not in A), and $A_1 \cup A_2 = B$. So by Axiom 3,

$$P(A_1) + P(A_2) = P(A_1 \cup A_2) = P(B).$$

In other words, $P(A) + P(B \setminus A) = P(B)$. Now $P(B \setminus A) \geq 0$ by Axiom 1; so

$$P(A) \leq P(B),$$

as we had to show.



Proposition 1.7

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

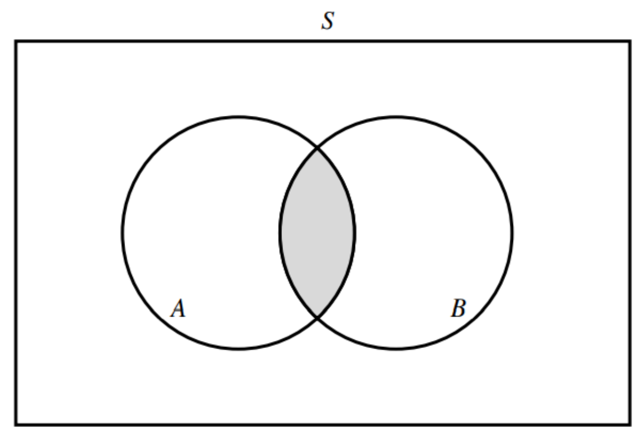
We now prove this from the axioms, using the Venn diagram as a guide. We see that $A \cup B$ is made up of three parts, namely

$$A_1 = A \cap B, \qquad A_2 = A \setminus B, \qquad A_3 = B \setminus A.$$

$$A \cup B = A_1 \cup A_2 \cup A_3$$

$$A_1 \cup A_2 = A$$

$$A_1 \cup A_3 = B$$



The sets A_1, A_2, A_3 are mutually disjoint. (We have three pairs of sets to check. Now $A_1 \cap A_2 = \emptyset$, since all elements of A_1 belong to B but no elements of A_2 do.

So, by Axiom 3, we have

$$\begin{aligned}
P(A) &= P(A_1) + P(A_2), \\
P(B) &= P(A_1) + P(A_3), \\
P(A \cup B) &= P(A_1) + P(A_2) + P(A_3).
\end{aligned}$$

From this we obtain

$$\begin{aligned}
P(A) + P(B) - P(A \cap B) &= (P(A_1) + P(A_2)) + (P(A_1) + P(A_3)) - P(A_1) \\
&= P(A_1) + P(A_2) + P(A_3) \\
&= P(A \cup B)
\end{aligned}$$

as required.

Proposition 1.8 *For any three events A, B, C, we have*

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).$$

Independence

Def: Two events A and B are said to be independent if:

$$P(A \cap B) = P(A) \cdot P(B).$$

Example: If we toss a coin more than once, or roll a die more than once, then you may assume that different tosses or rolls are independent.

More precisely, if we roll a fair six-sided die twice, then the probability of getting 4 on the first throw and 5 on the second is 1/36, since we assume that all 36 combinations of the two throws are equally likely. But (1/36) = (1/6)·(1/6), and the separate probabilities of getting 4 on the first throw and of getting 5 on the second are both equal to 1/6. So the two events are independent.

Independence

- Note: In general, it is always OK to assume that the outcomes of different tosses of a coin, or different throws of a die, are independent.
- Example: I have two red pens, one green pen, and one blue pen. I choose two pens without replacement. Let A be the event that I choose exactly one red pen, and B the event that I choose exactly one green pen. Is A and B are independent?

Sol: If the pens are called R1,R2,G,B, then:

S = {R1R2,R1G,R1B,R2G,R2B,GB},

A = {R1G,R1B,R2G,R2B},

B = {R1G,R2G,GB}

We have $P(A) = 4/6 = 2/3$, $P(B) = 3/6 = 1/2$, $P(A \cap B) = 2/6 = 1/3 = P(A)P(B)$, so A and B are independent.

Note: before you say ‘that’s obvious’, suppose that I have also a purple pen, and I do the same experiment. This time, if you write down the sample space and the two events and do the calculations, you will find that

$P(A) = 6/10 = 3/5$, $P(B) = 4/10 = 2/5$,

$P(A \cap B) = 2/10 = 1/5 \neq P(A)P(B)$.

That means, adding one more pen has made the events non-independent.

H.W. Consider the experiment where we toss a fair coin three times and note the results. Let A be the event ‘there are more heads than tails’, and B the event ‘the results of the first two tosses are the same’. Are A and B independent?

See you next Lecture

LECTURE 6#

Properties of independence

Proposition 1.11 *Let $A_1, \ldots, A_n$ be mutually independent. Then*

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \cdot P(A_2) \cdots P(A_n).$$

Proposition 1.12 *If A and B are independent, then A and B' are independent.*

We are given that $P(A \cap B) = P(A) \cdot P(B)$, and asked to prove that $P(A \cap B') = P(A) \cdot P(B')$.

From Corollary 4, we know that $P(B') = 1 - P(B)$. Also, the events $A \cap B$ and $A \cap B'$ are *disjoint* (since no outcome can be both in B and B'), and their union is A (since every event in A is either in B or in B'); so by Axiom 3, we have that $P(A) = P(A \cap B) + P(A \cap B')$. Thus,

$$\begin{aligned}
P(A \cap B') &= P(A) - P(A \cap B) \\
&= P(A) - P(A) \cdot P(B) \\
&\qquad \text{(since } A \text{ and } B \text{ are independent)} \\
&= P(A)(1 - P(B)) \\
&= P(A) \cdot P(B'),
\end{aligned}$$

which is what we were required to prove.

Independence

Corollary 1.13 *If A and B are independent, so are A' and B'.*

Apply the Proposition twice, first to A and B (to show that A and B' are independent), and then to B' and A (to show that B' and A' are independent).

Proposition 1.14 *Let events A, B, C be mutually independent. Then A and $B \cap C$ are independent, and A and $B \cup C$ are independent.*

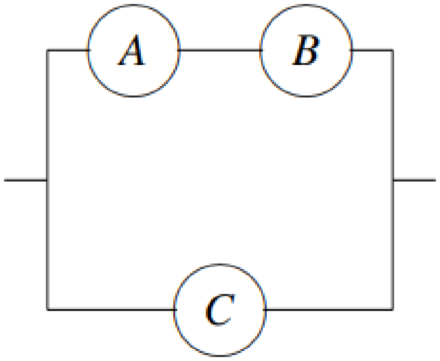
Example :The electrical apparatus in the diagram works so long as current can flow from left to right. The three components are independent. The probability that component A works is 0.8; the probability that component B works is 0.9; and the probability that component C works is 0.75. Find the probability that the apparatus works.

Independence

Sol: At risk of some confusion, we use the letters A, B and C for the events 'component A works', 'component B works', and 'component C works', respectively. Now the apparatus will work if either A and B are working, or C is working (or possibly both). Thus the event we are interested in is $(A \cap B) \cup C$.

Now

$$\begin{aligned} P((A\cap B)\cup C)) &= P(A\cap B)+P(C)-P(A\cap B\cap C) \\ &\qquad \text{(by Inclusion–Exclusion)} \\ &= P(A)\cdot P(B)+P(C)-P(A)\cdot P(B)\cdot P(C) \\ &\qquad \text{(by mutual independence)} \\ &= (0.8)\cdot(0.9)+(0.75)-(0.8)\cdot(0.9)\cdot(0.75) \\ &= 0.93. \end{aligned}$$



Conditional probability

The *conditional probability of an event* A, given that an event B has occurred, is equal to

$$P(A|B) = \frac{P(A \cap B)}{P(B)},$$

provided $P(B) > 0$. [The symbol $P(A|B)$ is read "probability of A given B."]

Ex: Suppose that a balanced die is tossed once. Find the probability of a 1, given that an odd number was obtained

Solution Define these events:

A: Observe a 1.
B: Observe an odd number.

We seek the probability of A given that the event B has occurred. The event $A \cap B$ requires the observance of both a 1 and an odd number. In this instance, $A \subset B$, so $A \cap B = A$ and $P(A \cap B) = P(A) = 1/6$. Also, $P(B) = 1/2$ and, using Definition 2.9,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{1/2} = \frac{1}{3}.$$

Conditional probability

Note :There is a connection between conditional probability and independence:

Proposition 2.1 *Let A and B be events with $P(B) \neq 0$. Then A and B are independent if and only if $P(A \mid B) = P(A)$.*

Proof The words 'if and only if' tell us that we have two jobs to do: we have to show that if A and B are independent, then $P(A \mid B) = P(A)$; and that if $P(A \mid B) = P(A)$, then A and B are independent.

So first suppose that A and B are independent. Remember that this means that $P(A \cap B) = P(A) \cdot P(B)$. Then

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B)}{P(B)} = P(A),$$

that is, $P(A \mid B) = P(A)$, as we had to prove.

Now suppose that $P(A \mid B) = P(A)$. In other words,

$$\frac{P(A \cap B)}{P(B)} = P(A),$$

using the definition of conditional probability. Now clearing fractions gives

$$P(A \cap B) = P(A) \cdot P(B),$$

which is just what the statement 'A and B are independent' means.

Ex: Consider the following events in the toss of a single die:

A: Observe an odd number.

B: Observe an even number.

C: Observe a 1 or 2.

a) Are A and B independent events?

b) Are A and C independent events?

Sol:

a) To decide whether A and B are independent, we must see whether they satisfy the conditions of Proposition 2.1 . In this example, $P(A) = 1/2$, $P(B) = 1/2$, and $P(C) = 1/3$. Because $A \cap B = \emptyset$, $P(A|B) = 0$, and it is clear that $P(A|B) \neq P(A)$. Events A and B are dependent events.

b)Are A and C independent? Note that $P(A|C) = 1/2$ and, as before, $P(A) = 1/2$. Therefore, $P(A|C) = P(A)$, and A and C are independent

H.W: Three brands of coffee, X, Y , and Z, are to be ranked according to taste by a judge. Define the following events:

A: Brand X is preferred to Y .

B: Brand X is ranked best.

C: Brand X is ranked second best.

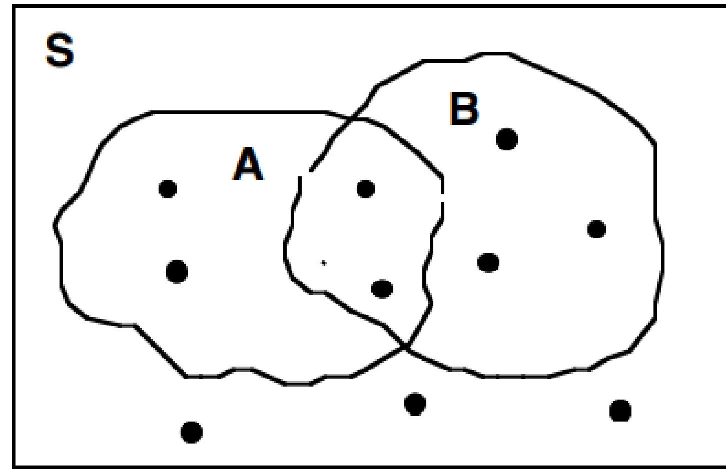
D: Brand X is ranked third best.

If the judge actually has no taste preference and randomly assigns ranks to the brands, is event A independent of events B, C, and D?

See you next Lecture

LECTURE 7#

Example. The following diagram shows two events A and B in the sample space S. Are the events A and B independent?



Answer: There are 10 black dots in S and event A contains 4 of these dots. So the probability of A, is P(A) $=\frac{4}{10}$. Similarly, event B contains 5 black dots. Hence P(B) $= \frac{5}{10}$. The conditional probability of A given B is

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{2}{5}.$$

This shows that P(A|B) = P(A). Hence A and B are independent.

Two Laws of Probability

The following law give the probabilities of unions of events.

THEOREM 1: (The Multiplicative Law of Probability)

The probability of the intersection of two events A and B is

$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B).$$

If A and B are independent, then $P(A \cap B) = P(A)P(B)$

Proof: H.W.

Theorem 2: (Theorem of Total Probability)

Let A1,A2,...,An form a partition of the sample space with $P(Ai) \neq 0$ for all i, and let B be any event. Then

$$P(B) = \sum_{i=1}^{n} P(B \mid A_i) \cdot P(A_i).$$

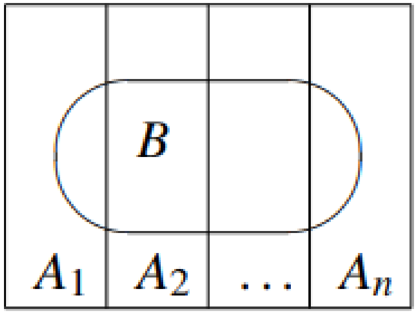
Proof By definition, $P(B \mid A_i) = P(B \cap A_i)/P(A_i)$. Multiplying up, we find that

$$P(B \cap A_i) = P(B \mid A_i) \cdot P(A_i).$$

Now consider the events $B \cap A_1, B \cap A_2, \ldots, B \cap A_n$. These events are pairwise disjoint; for any outcome lying in both $B \cap A_i$ and $B \cap A_j$ would lie in both A_i and A_j, and by assumption there are no such outcomes. Moreover, the union of all these events is B, since every outcome lies in one of the A_i. So, by Axiom 3, we conclude that

$$\sum_{i=1}^{n} P(B \cap A_i) = P(B).$$

Substituting our expression for $P(B \cap A_i)$ gives the result.



Example 1: An ice-cream seller has to decide whether to order more stock for the Bank Holiday weekend. He estimates that, if the weather is sunny, he has a 90% chance of selling all his stock; if it is cloudy, his chance is 60%; and if it rains, his chance is only 20%. According to the weather forecast, the probability of sunshine is 30%, the probability of cloud is 45%, and the probability of rain is 25%. (We assume that these are all the possible outcomes, so that their probabilities must add up to 100%.) What is the overall probability that the salesman will sell all his stock?

Sol: Let A1 be the event 'it is sunny', A2 the event 'it is cloudy', and A3 the event 'it is rainy'. Then A1, A2 and A3 form a partition of the sample space, and we are given that:

$$P(A1) = 0.3,\ P(A2) = 0.45,\ P(A3) = 0.25.$$

Let B be the event 'the salesman sells all his stock'. The other information we are given is that $P(B \mid A1) = 0.9$, $P(B \mid A2) = 0.6$, $P(B \mid A3) = 0.2$. By the Theorem of Total Probability, $P(B) = (0.9\times0.3) + (0.6\times0.45) + (0.2\times0.25) = 0.59$.

Bayes' Theorem

Theorem : Let A and B be events with non-zero probability. Then

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

Proof: $P(A \mid B) \cdot P(B) = P(A \cap B) = P(B \mid A) \cdot P(A)$,

Ex : For same Example 1, we are asked for $P(A1 \mid B)$. We were given that $P(B \mid A1) = 0.9$ and that $P(A1) = 0.3$, and we calculated that $P(B) = 0.59$. So by Bayes' Theorem,

$$P(A_1 \mid B) = \frac{P(B \mid A_1)P(A_1)}{P(B)} = \frac{0.9 \times 0.3}{0.59} = 0.46$$

Conditional probability

Example 2: 2% of the population have a certain blood disease in a serious form; 10% have it in a mild form; and 88% don't have it at all. A new blood test is developed; the probability of testing positive is 9/10 if the subject has the serious form, 6/10 if the subject has the mild form, and 1/10 if the subject doesn't have the disease. I have just tested positive. What is the probability that I have the serious form of the disease?

Sol:

Let A_1 be 'has disease in serious form', A_2 be 'has disease in mild form', and A_3 be 'doesn't have disease'. Let B be 'test positive'. Then we are given that A_1, A_2, A_3 form a partition and

$$\begin{array}{lll} P(A_1) = 0.02 & P(A_2) = 0.1 & P(A_3) = 0.88 \\ P(B \mid A_1) = 0.9 & P(B \mid A_2) = 0.6 & P(B \mid A_3) = 0.1 \end{array}$$

Thus, by the Theorem of Total Probability,

$$P(B) = 0.9 \times 0.02 + 0.6 \times 0.1 + 0.1 \times 0.88 = 0.166,$$

and then by Bayes' Theorem,

$$P(A_1 \mid B) = \frac{P(B \mid A_1)P(A_1)}{P(B)} = \frac{0.9 \times 0.02}{0.166} = 0.108$$

See you next Lecture

LECTURE 8#

The cumulative distribution function

Definition 1. The cumulative distribution function(CDF) F(x) of a random variable X is defined as

$$F(x) = P(X \leq x) \qquad \text{for all real numbers x}$$

Theorem 1. If X is a random variable with the space R_X, then

$F(x) = \sum_{t \leq x} p(t)$ for all $t \in R_X$.

Example 1. If the probability density function of the random variable X is given by

$$\frac{1}{144}(2x-1) \qquad \text{for } x = 1, 2, 3, ..., 12$$

then find the cumulative distribution function of X.

Answer: The space of the random variable X is given by $R_X = \{1,\ 2,\ 3, ...,\ 12\}$.

Then

$$F(1) = \sum_{t \le 1} f(t) = f(1) = \frac{1}{144}$$

$$F(2) = \sum_{t \le 2} f(t) = f(1) + f(2) = \frac{1}{144} + \frac{3}{144} = \frac{4}{144}$$

$$F(3) = \sum_{t \le 3} f(t) = f(1) + f(2) + f(3) = \frac{1}{144} + \frac{3}{144} + \frac{5}{144} = \frac{9}{144}$$

.. ..

.. ..

$$F(12) = \sum_{t \le 12} f(t) = f(1) + f(2) + \cdots + f(12) = 1.$$

Theorem 2. Let X be a random variable with cumulative distribution function F(x). Then the cumulative distribution function satisfies the followings

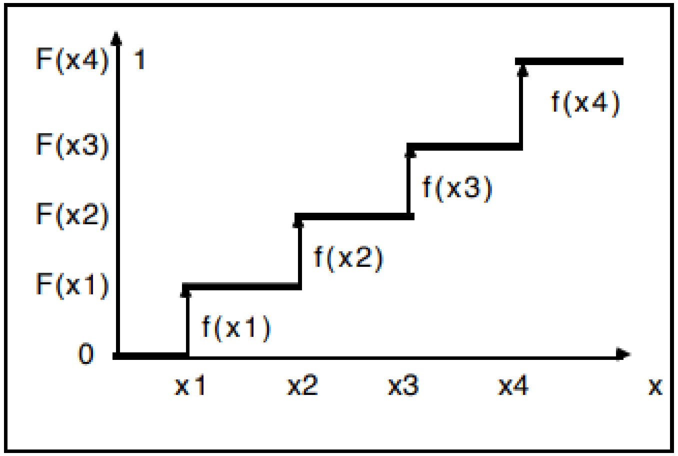
(a) $F(-\infty) = 0$,

(b) $F(\infty) = 1$, and

(c) $F(x)$ is an increasing function, that is if $x < y$, then $F(x) \le F(y)$ for all reals x, y.

Theorem 3 If the space R_X of the random variable X is given by $R_X = \{x_1 < x_2 < x_3 < \cdots < x_n\}$ then

$$
\begin{aligned}
f(x_1) &= F(x_1) \\
f(x_2) &= F(x_2) - F(x_1) \\
f(x_3) &= F(x_3) - F(x_2) \\
&\;\;\cdots\cdots\cdots \\
&\;\;\cdots\cdots\cdots \\
f(x_n) &= F(x_n) - F(x_{n-1}).
\end{aligned}
$$



Theorem 1 tells us how to find cumulative distribution function from the probability density function, whereas Theorem 2 tells us how to find the probability density function given the cumulative distribution function.

Example 2. Find the probability density function of the random variable X whose cumulative distribution function is

$$F(x) = \begin{cases} 0.00 & \text{if } x < -1 \\ 0.25 & \text{if } -1 \leq x < 1 \\ 0.50 & \text{if } 1 \leq x < 3 \\ 0.75 & \text{if } 3 \leq x < 5 \\ 1.00 & \text{if } x \geq 5 \;. \end{cases}$$

Also, find (a) $P(X \leq 3)$, (b) $P(X = 3)$, and (c) $P(X < 3)$.

Answer: The space of this random variable is given by

$$R_X = \{-1,\, 1,\, 3,\, 5\}.$$

By the previous theorem, the probability density function of X is given by

$$\begin{aligned} f(-1) &= 0.25 \\ f(1) &= 0.50 - 0.25 = 0.25 \\ f(3) &= 0.75 - 0.50 = 0.25 \\ f(5) &= 1.00 - 0.75 = 0.25. \end{aligned}$$

The probability $P(X \leq 3)$ can be computed by using the definition of F. Hence

$$P(X \leq 3) = F(3) = 0.75.$$

The probability P(X = 3) can be computed from
P(X = 3) = F(5) − F(3) = 1 − 0.75 = 0.25.
Finally, we get P(X < 3) from P(X < 3) = P(X ≤ 1) = F(1) = 0.5.

Moments of Random Variables

Definition 4.1. The n^{th} moment about the origin of a random variable X, as denoted by $E(X^n)$, is defined to be

$$E\left(X^n\right) = \begin{cases} \sum\limits_{x \in R_X} x^n f(x) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} x^n f(x)\, dx & \text{if } X \text{ is continuous} \end{cases}$$

- If n = 1, then E(X) is called the first moment about the origin.

- If n = 2, then E(X^2) is called the second moment of X about the origin
- In general, these moments may or may not exist for a given random variable.

Expected Value of Random Variables

Definition 4.2. Let X be a random variable with space R_X and probability density function $f(x)$. The mean μ_X of the random variable X is defined as

$$\mu_X = \begin{cases} \displaystyle\sum_{x \in R_X} x\, f(x) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} x\, f(x)\, dx & \text{if } X \text{ is continuous} \end{cases}$$

if the right hand side exists.

The mean of a random variable is a composite of its values weighted by the corresponding probabilities. The mean is a measure of central tendency: the value that the random variable takes "on average." The mean is also called the expected value of the random variable X and is denoted by E(X). The symbol E is called the expectation operator. The expected value of a random variable may or may not exist.

Let Y be a discrete random variable with the probability function p(y). Then the expected value of Y , E(Y), is defined to be $E(Y) = \sum yp(y)$.

Example: The probability distribution for a random variable Y is given in Table 3.3. Find the mean.

Table 3.3 Probability distribution for Y

y	$p(y)$
0	1/8
1	1/4
2	3/8
3	1/4

Solution

$$\mu = E(Y) = \sum_{y=0}^{3} yp(y) = (0)(1/8) + (1)(1/4) + (2)(3/8) + (3)(1/4) = 1.75,$$

x	0	1	2	3
$f(x)$	0.01	0.32	0.46	0.21

x	0	1	2	3
$p(x)$	0.1	0.32	0.46	0.21

x	0	1	2	3
$f(x)$	0.01	-0.32	0.46	0.21

$\sum P(x) = P(0) +$

$P(1) + \quad = 1$

See you next Lecture

LECTURE 9#

Example 1 In how many ways a committee consisting of 3 men and 2 women, can be chosen from 7 men and 5 women?

Example 2 A group consists of 4 girls and 7 boys. In how many ways can a team of 5 members be selected if the team has

(i) no girls

(ii) at least one boy and one girl

(iii) at least three girls.

Example 3 In how many ways can a supermarket manager display 5 brands of cereals in 3 spaces on a shelf?

Example 4 There are 15 balls numbered 1 to 15, in a bag. If a person selects one at random, what is the probability that the number printed on the ball will be a prime number greater than 5?

Example 5: Let A denote the event 'student is female' and let B denote the event 'student is French'. In a class of 100 students suppose 60 are French, and suppose that 10 of the French students are females. Find the probability that if I pick a French student, it will be a girl, that is, find $P(A|B)$

Example 6: What is the probability that the total of two dice will be greater than 8, given that the first die is a 6?

Example 7 It is known that the probability of obtaining zero defectives in a sample of 40 items is 0.34 while the probability of obtaining 1 defective item in the sample is 0.46. What is the probability of

(a) obtaining not more than 1 defective item in a sample?

(b) obtaining more than 1 defective items in a sample?

See you next Lecture

LECTURE 10#

continuous random variable

Definition 3.7. Let X be a continuous random variable whose space is the set of real numbers $\mathbb{R}$. A nonnegative real valued function $f : \mathbb{R} \to \mathbb{R}$ is said to be the probability density function for the continuous random variable X if it satisfies:

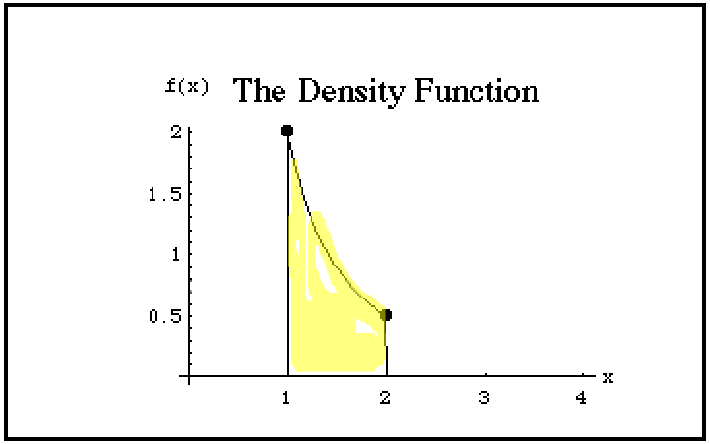
(a) $\int_{-\infty}^{\infty} f(x)\, dx = 1$, and

(b) if A is an event, then $P(A) = \int_A f(x)\, dx$.

Example 3.10. Is the real valued function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 2\,x^{-2} & \text{if } 1 < x < 2 \\ 0 & \text{otherwise,} \end{cases}$$

a probability density function for some random variable X?



Answer: We have to show that f is nonnegative and the area under $f(x)$ is unity. Since the domain of f is the interval $(0, 1)$, it is clear that f is nonnegative. Next, we calculate

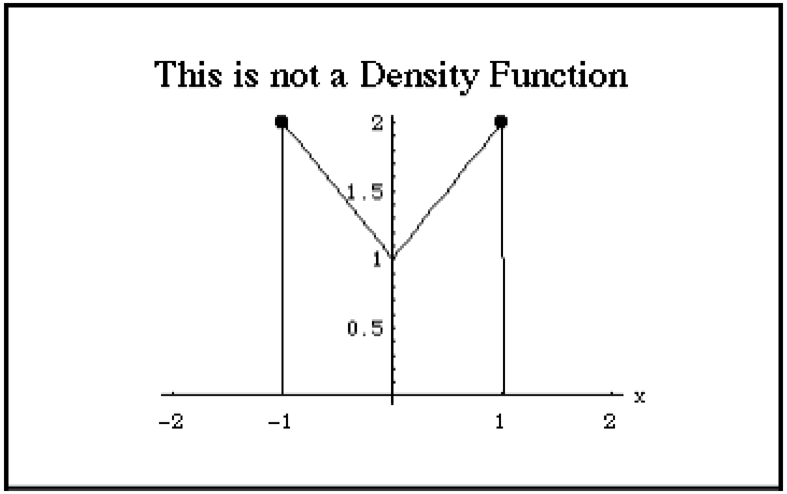
$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)\,dx &= \int_{1}^{2} 2\,x^{-2}\,dx \\
&= -2\left[\frac{1}{x}\right]_1^2 \\
&= -2\left[\frac{1}{2} - 1\right] \\
&= 1.
\end{aligned}$$

Thus f is a probability density function.

Example 3.11. Is the real valued function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1 + |x| & \text{if } -1 < x < 1 \\ 0 & \text{otherwise,} \end{cases}$$

a probability density function for some random variable X?



Answer: It is easy to see that f is nonnegative, that is $f(x) \geq 0$, since $f(x) = 1 + |x|$. Next we show that the area under f is not unity. For this we compute

$$\begin{aligned}
\int_{-1}^{1} f(x)\,dx &= \int_{-1}^{1} (1 + |x|)\,dx \\
&= \int_{-1}^{0} (1 - x)\,dx + \int_{0}^{1} (1 + x)\,dx \\
&= \left[x - \frac{1}{2}x^2\right]_{-1}^{0} + \left[x + \frac{1}{2}x^2\right]_{0}^{1} \\
&= 1 + \frac{1}{2} + 1 + \frac{1}{2} \\
&= 3.
\end{aligned}$$

Thus f is not a probability density function for some random variable X.

Example 3.12. For what value of the constant c, the real valued function $f : \mathbb{R} \to \mathbb{R}$ given by

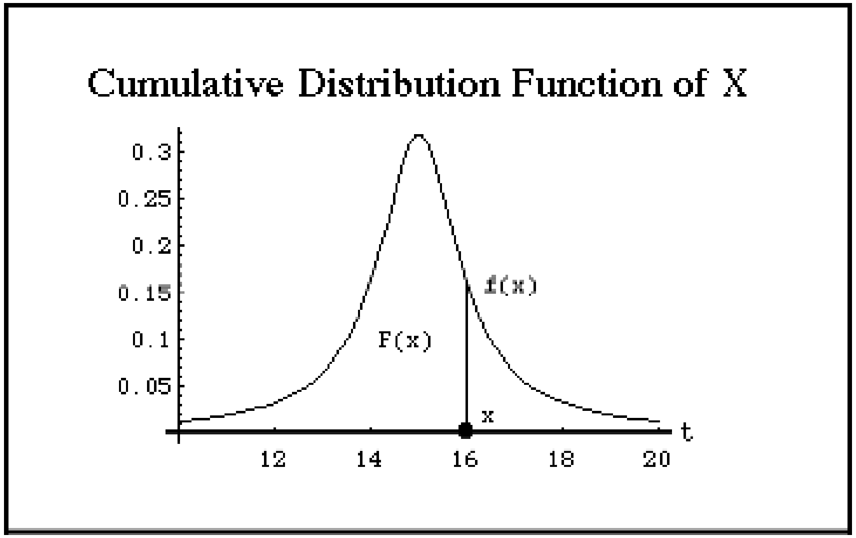
$$f(x) = \frac{c}{1 + (x - \theta)^2}, \qquad -\infty < x < \infty,$$

where θ is a real parameter, is a probability density function for random variable X?

Definition 3.8. Let $f(x)$ be the probability density function of a continuous random variable X. The cumulative distribution function $F(x)$ of X is defined as

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt.$$

The cumulative distribution function $F(x)$ represents the area under the probability density function $f(x)$ on the interval $(-\infty, x)$ (see figure below).



Like the discrete case, the cdf is an increasing function of x, and it takes value 0 at negative infinity and 1 at positive infinity.

Theorem 3.5. If $F(x)$ is the cumulative distribution function of a continuous random variable X, the probability density function $f(x)$ of X is the derivative of $F(x)$, that is

$$\frac{d}{dx}F(x) = f(x).$$

Proof: By Fundamental Theorem of Calculus, we get

$$\begin{aligned} \frac{d}{dx}\left(F(x)\right) &= \frac{d}{dx}\left(\int_{-\infty}^{x} f(t)\,dt\right) \\ &= f(x)\frac{dx}{dx} \\ &= f(x). \end{aligned}$$

This theorem tells us that if the random variable is continuous, then we can find the pdf given cdf by taking the derivative of the cdf. Recall that for discrete random variable

Example 3.15. What is the probability density function of the random variable whose cdf is

$$F(x) = \frac{1}{1+e^{-x}}, \qquad -\infty < x < \infty?$$

Answer: The pdf of the random variable is given by

$$\begin{aligned} f(x) &= \frac{d}{dx} F(x) \\ &= \frac{d}{dx}\left(\frac{1}{1+e^{-x}}\right) \\ &= \frac{d}{dx}\left(1+e^{-x}\right)^{-1} \\ &= (-1)\left(1+e^{-x}\right)^{-2} \frac{d}{dx}\left(1+e^{-x}\right) \\ &= \frac{e^{-x}}{(1+e^{-x})^2}. \end{aligned}$$

Theorem 3.6. Let X be a continuous random variable whose cdf is $F(x)$. Then followings are true:

(a) $P(X < x) = F(x)$,

(b) $P(X > x) = 1 - F(x)$,

(c) $P(X = x) = 0$, and

(d) $P(a < X < b) = F(b) - F(a)$.

EXAMPLE 3.12 : (a) Find the constant c such that the function $f(x) = \begin{cases} cx^2 & 0 < x < 3 \\ 0 & \text{otherwise} \end{cases}$ is a density function.

(b) compute P(1 < X <2)

(c) Find the distribution function

(d) Use the result of (c) to find P(1 < x ≤ 2) .

Solution:

(a) Since $f(x)$ satisfies Property 1 if $c \geq 0$, it must satisfy Property 2 in order to be a density function. Now

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^3 cx^2\,dx = \left.\frac{cx^3}{3}\right|_0^3 = 9c$$

and since this must equal 1, we have $c = 1/9$.

(b)
$$P(1 < X < 2) = \int_1^2 \frac{1}{9}x^2\,dx = \left.\frac{x^3}{27}\right|_1^2 = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}$$

(c) We have $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du$

If $x < 0$, then $F(x) = 0$. If $0 \leq x < 3$, then

$$F(x) = \int_0^x f(u)\,du = \int_0^x \frac{1}{9}u^2\,du = \frac{x^3}{27}$$

If $x \geq 3$, then

$$F(x) = \int_0^3 f(u)\,du + \int_3^x f(u)\,du = \int_0^3 \frac{1}{9}u^2\,du + \int_3^x 0\,du = 1$$

Thus the required distribution function is

$$F(x) = \begin{cases} 0 & x < 0 \\ x^3/27 & 0 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

(d) We have

$$\begin{aligned} P(1 < X \leq 2) &= P(X \leq 2) - P(X \leq 1) \\ &= F(2) - F(1) \\ &= \frac{2^3}{27} - \frac{1^3}{27} = \frac{7}{27} \end{aligned}$$

H.W.

1) Let Y possess a density function

$$f(y) = \begin{cases} c(2-y), & 0 \le y \le 2, \\ 0, & \text{elsewhere.} \end{cases}$$

a Find c.

b Find $F(y)$.

c Graph $f(y)$ and $F(y)$.

d Use $F(y)$ in part (b) to find $P(1 \le Y \le 2)$.

2) Let Y have the density function given by

$$f(y) = \begin{cases} .2, & -1 < y \le 0, \\ .2 + cy, & 0 < y \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

a Find c.

b Find $F(y)$.

c Graph $f(y)$ and $F(y)$.

d Use $F(y)$ in part (b) to find $F(-1)$, $F(0)$, and $F(1)$.

e Find $P(0 \le Y \le .5)$.

See you next Lecture

LECTURE 11#

Examples

1) Suppose that

$$F(y) = \begin{cases} 0, & \text{for } y < 0, \\ y, & \text{for } 0 \le y \le 1, \\ 1, & \text{for } y > 1. \end{cases}$$

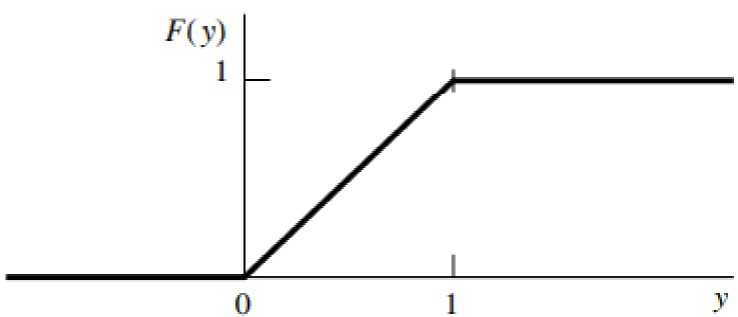
Find the probability density function for Y and graph it.

Solution Because the density function $f(y)$ is the derivative of the distribution function $F(y)$, when the derivative exists,

$$f(y) = \frac{dF(y)}{dy} = \begin{cases} \dfrac{d(0)}{dy} = 0, & \text{for } y < 0, \\ \dfrac{d(y)}{dy} = 1, & \text{for } 0 < y < 1, \\ \dfrac{d(1)}{dy} = 0, & \text{for } y > 1, \end{cases}$$

and $f(y)$ is undefined at $y = 0$ and $y = 1$. A graph of $F(y)$ is shown in Figure 4.4.

FIGURE **4.4**
Distribution function $F(y)$ for Example 4.2



■

2)

Let Y be a continuous random variable with probability density function given by

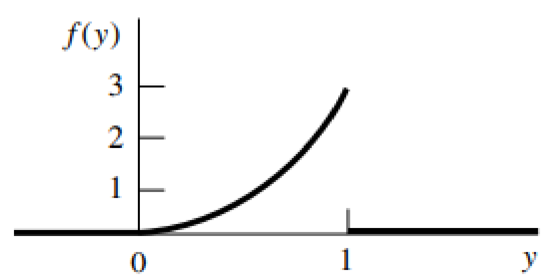
$$f(y) = \begin{cases} 3y^2, & 0 \le y \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Find $F(y)$. Graph both $f(y)$ and $F(y)$.

$$F(y) = \int_{-\infty}^{y} f(t)\,dt,$$

we have, for this example,

$$F(y) = \begin{cases} \int_{-\infty}^{y} 0\,dt = 0, & \text{for } y < 0, \\ \int_{-\infty}^{0} 0\,dt + \int_{0}^{y} 3t^2\,dt = 0 + t^3\big]_0^y = y^3, & \text{for } 0 \le y \le 1, \\ \int_{-\infty}^{0} 0\,dt + \int_{0}^{1} 3t^2\,dt + \int_{1}^{y} 0\,dt = 0 + t^3\big]_0^1 + 0 = 1, & \text{for } 1 < y. \end{cases}$$



3)

Given $f(y) = cy^2, 0 \le y \le 2$, and $f(y) = 0$ elsewhere, find the value of c for which $f(y)$ is a valid density function.

4)

Let the distribution function of a random variable Y be

$$F(y) = \begin{cases} 0, & y \le 0, \\ \frac{y}{8}, & 0 < y < 2, \\ \frac{y^2}{16}, & 2 \le y < 4, \\ 1, & y \ge 4. \end{cases}$$

a Find the density function of Y.

b Find $P(1 \le Y \le 3)$.

c Find $P(Y \ge 1.5)$.

d Find $P(Y \ge 1 | Y \le 3)$.

Expected value and variance

Definition : Let X be a random variable with space S and probability density function f(x). The mean μ (expected value) of the random variable X is defined as

$$\mu_X = \begin{cases} \sum_{x \in R_X} x\, f(x) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} x\, f(x)\, dx & \text{if } X \text{ is continuous} \end{cases}$$

Theorem :Let g(Y) be a function of Y ; then the expected value of g(Y) is given by

$$E[g(Y)] = \int_{-\infty}^{\infty} g(y) f(y)\, dy,$$

Theorem

Let c be a constant and let $g(Y)$, $g_1(Y)$, $g_2(Y)$, ..., $g_k(Y)$ be functions of a continuous random variable Y. Then the following results hold:

1. $E(c) = c$.
2. $E[cg(Y)] = cE[g(Y)]$.
3. $E[g_1(Y)+g_2(Y)+\cdots+g_k(Y)] = E[g_1(Y)]+E[g_2(Y)]+\cdots+E[g_k(Y)]$.

Expected value and variance

Definition : The variance of X is the number Var(X) given by

$$\mathrm{Var}(X) = E(X^2) - E(X)^2.$$

Theorem 4.2. If X is a random variable with mean μ_X and variance σ_X^2, then

$$\sigma_X^2 = E(X^2) - (\mu_X)^2.$$

Proof:

$$\begin{aligned}
\sigma_X^2 &= E\left([X - \mu_X]^2 \right) \\
&= E\left(X^2 - 2\mu_X X + \mu_X^2 \right) \\
&= E(X^2) - 2\mu_X E(X) + (\mu_X)^2 \\
&= E(X^2) - 2\mu_X \mu_X + (\mu_X)^2 \\
&= E(X^2) - (\mu_X)^2.
\end{aligned}$$

Example:

we determined that $f(y) = (3/8)y^2$ for $0 \le y \le 2$, $f(y) = 0$ elsewhere

find $\mu = E(Y)$ and $\sigma^2 = V(Y)$.

Solution:

$$E(Y) = \int_{-\infty}^{\infty} y f(y)\, dy$$
$$= \int_0^2 y \left(\frac{3}{8}\right) y^2\, dy$$
$$= \left(\frac{3}{8}\right)\left(\frac{1}{4}\right) y^4 \Big]_0^2 = 1.5.$$

The variance of Y can be found once we determine $E(Y^2)$. In this case,

$$E(Y^2) = \int_{-\infty}^{\infty} y^2 f(y)\, dy$$
$$= \int_0^2 y^2 \left(\frac{3}{8}\right) y^2\, dy$$
$$= \left(\frac{3}{8}\right)\left(\frac{1}{5}\right) y^5 \Big]_0^2 = 2.4.$$

Thus, $\sigma^2 = V(Y) = E(Y^2) - [E(Y)]^2 = 2.4 - (1.5)^2 = 0.15$.

Theorem 4.3. If X is a random variable with mean μ_X and variance σ_X^2, then

$$Var(aX + b) = a^2\, Var(X),$$

where a and b are arbitrary real constants.

Proof:

$$\begin{aligned}
Var(a\,X + b) &= E\left(\left[\,(a\,X + b) - \mu_{a\,X+b}\,\right]^2\right)\\
&= E\left(\left[\,a\,X + b - E(a\,X + b)\,\right]^2\right)\\
&= E\left(\left[\,a\,X + b - a\,\mu_{X+} - b\right]^2\right)\\
&= E\left(a^2\,\left[\,X - \mu_X\,\right]^2\right)\\
&= a^2\,E\left(\left[\,X - \mu_X\,\right]^2\right)\\
&= a^2\,Var(X).
\end{aligned}$$

Example. Let X have the density function

$$f(x) = \begin{cases} \frac{2x}{k^2} & \text{for } 0 \leq x \leq k \\ 0 & \text{otherwise.} \end{cases}$$

For what value of k is the variance of X equal to 2?

Answer: The expected value of X is

$$\begin{aligned} E(X) &= \int_0^k x\, f(x)\, dx \\ &= \int_0^k x\, \frac{2x}{k^2}\, dx \\ &= \frac{2}{3}\, k. \end{aligned}$$

$$\begin{aligned} E(X^2) &= \int_0^k x^2\, f(x)\, dx \\ &= \int_0^k x^2\, \frac{2x}{k^2}\, dx \\ &= \frac{2}{4}\, k^2. \end{aligned}$$

Hence, the variance is given by

$$\begin{aligned} Var(X) &= E(X^2) - (\mu_X)^2 \\ &= \frac{2}{4}\, k^2 - \frac{4}{9}\, k^2 \\ &= \frac{1}{18}\, k^2. \end{aligned}$$

Since this variance is given to be 2, we get

$$\frac{1}{18}\, k^2 = 2$$

and this implies that $k = \pm 6$. But k is given to be greater than 0, hence k must be equal to 6.

Example: If the probability density function of the random variable is

then what is the variance of X?

$$f(x) = \begin{cases} 1 - |x| & \text{for } |x| < 1 \\ 0 & \text{otherwise,} \end{cases}$$

Answer: Since $Var(X) = E(X^2) - \mu_X^2$, we need to find the first and second moments of X. The first moment of X is given by

$$\begin{aligned} \mu_X &= E(X) \\ &= \int_{-\infty}^{\infty} x\, f(x)\, dx \\ &= \int_{-1}^{1} x\,(1 - |x|)\, dx \\ &= \int_{-1}^{0} x\,(1 + x)\, dx + \int_{0}^{1} x\,(1 - x)\, dx \\ &= \int_{-1}^{0} (x + x^2)\, dx + \int_{0}^{1} (x - x^2)\, dx \\ &= \frac{1}{3} - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} \\ &= 0. \end{aligned}$$

The second moment $E(X^2)$ of X is given by

$$\begin{aligned}
E(X^2) &= \int_{-\infty}^{\infty} x^2 \, f(x) \, dx \\
&= \int_{-1}^{1} x^2 \, (1 - |x|) \, dx \\
&= \int_{-1}^{0} x^2 \, (1 + x) \, dx + \int_{0}^{1} x^2 \, (1 - x) \, dx \\
&= \int_{-1}^{0} (x^2 + x^3) \, dx + \int_{0}^{1} (x^2 - x^3) \, dx \\
&= \frac{1}{3} - \frac{1}{4} + \frac{1}{3} - \frac{1}{4} \\
&= \frac{1}{6}.
\end{aligned}$$

Thus, the variance of X is given by

$$Var(X) = E(X^2) - \mu_X^2 = \frac{1}{6} - 0 = \frac{1}{6}.$$

See you next Lecture

LECTURE 12#

Expected value and variance

Definition : Let X be a random variable with space S and probability density function f(x). The mean μ (expected value) of the random variable X is defined as

$$\mu_X = \begin{cases} \sum_{x \in R_X} x\, f(x) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} x\, f(x)\, dx & \text{if } X \text{ is continuous} \end{cases}$$

Theorem :Let g(Y) be a function of Y ; then the expected value of g(Y) is given by

$$E\,[g(Y)] = \int_{-\infty}^{\infty} g(y) f(y)\, dy,$$

Theorem

Let c be a constant and let $g(Y)$, $g_1(Y)$, $g_2(Y)$, ..., $g_k(Y)$ be functions of a continuous random variable Y. Then the following results hold:

1. $E(c) = c$.
2. $E[cg(Y)] = cE[g(Y)]$.
3. $E[g_1(Y)+g_2(Y)+\cdots+g_k(Y)] = E[g_1(Y)]+E[g_2(Y)]+\cdots+E[g_k(Y)]$.

Expected value and variance

Definition : The variance of X is the number Var(X) given by

$$\mathrm{Var}(X) = E(X^2) - E(X)^2.$$

Theorem 4.2. If X is a random variable with mean μ_X and variance σ_X^2, then

$$\sigma_X^2 = E(X^2) - (\mu_X)^2.$$

Proof:

$$\begin{aligned}
\sigma_X^2 &= E\left([X - \mu_X]^2\right) \\
&= E\left(X^2 - 2\mu_X X + \mu_X^2\right) \\
&= E(X^2) - 2\mu_X E(X) + (\mu_X)^2 \\
&= E(X^2) - 2\mu_X \mu_X + (\mu_X)^2 \\
&= E(X^2) - (\mu_X)^2.
\end{aligned}$$

Example:

we determined that $f(y) = (3/8)y^2$ for $0 \leq y \leq 2$, $f(y) = 0$ elsewhere

find $\mu = E(Y)$ and $\sigma^2 = V(Y)$.

Solution:

$$E(Y) = \int_{-\infty}^{\infty} y f(y)\, dy$$
$$= \int_0^2 y \left(\frac{3}{8}\right) y^2\, dy$$
$$= \left(\frac{3}{8}\right)\left(\frac{1}{4}\right) y^4\Big]_0^2 = 1.5.$$

The variance of Y can be found once we determine $E(Y^2)$. In this case,

$$E(Y^2) = \int_{-\infty}^{\infty} y^2 f(y)\, dy$$
$$= \int_0^2 y^2 \left(\frac{3}{8}\right) y^2\, dy$$
$$= \left(\frac{3}{8}\right)\left(\frac{1}{5}\right) y^5\Big]_0^2 = 2.4.$$

Thus, $\sigma^2 = V(Y) = E(Y^2) - [E(Y)]^2 = 2.4 - (1.5)^2 = 0.15.$

Theorem 4.3. If X is a random variable with mean μ_X and variance σ_X^2, then

$$Var(aX+b) = a^2\,Var(X),$$

where a and b are arbitrary real constants.

Proof:

$$\begin{aligned}
Var(a\,X+b) &= E\left(\left[\,(a\,X+b) - \mu_{aX+b}\,\right]^2\right) \\
&= E\left(\left[\,a\,X+b - E(a\,X+b)\,\right]^2\right) \\
&= E\left(\left[\,a\,X+b - a\,\mu_{X+} - b\right]^2\right) \\
&= E\left(a^2\,\left[\,X-\mu_X\,\right]^2\right) \\
&= a^2\,E\left(\left[\,X-\mu_X\,\right]^2\right) \\
&= a^2\,Var(X).
\end{aligned}$$

Example. Let X have the density function

$$f(x) = \begin{cases} \frac{2x}{k^2} & \text{for } 0 \leq x \leq k \\ 0 & \text{otherwise.} \end{cases}$$

For what value of k is the variance of X equal to 2?

Answer: The expected value of X is

$$E(X) = \int_0^k x\, f(x)\, dx = \int_0^k x\, \frac{2x}{k^2}\, dx = \frac{2}{3}\, k.$$

$$E(X^2) = \int_0^k x^2\, f(x)\, dx = \int_0^k x^2\, \frac{2x}{k^2}\, dx = \frac{2}{4}\, k^2.$$

Hence, the variance is given by

$$Var(X) = E(X^2) - (\mu_X)^2 = \frac{2}{4}\, k^2 - \frac{4}{9}\, k^2 = \frac{1}{18}\, k^2.$$

Since this variance is given to be 2, we get

$$\frac{1}{18}\, k^2 = 2$$

and this implies that $k = \pm 6$. But k is given to be greater than 0, hence k must be equal to 6.

Example: If the probability density function of the random variable is

$$f(x) = \begin{cases} 1 - |x| & \text{for } |x| < 1 \\ 0 & \text{otherwise,} \end{cases}$$

then what is the variance of X?

Answer: Since $Var(X) = E(X^2) - \mu_X^2$, we need to find the first and second moments of X. The first moment of X is given by

$$\begin{aligned}
\mu_X &= E(X) \\
&= \int_{-\infty}^{\infty} x\, f(x)\, dx \\
&= \int_{-1}^{1} x\,(1 - |x|)\, dx \\
&= \int_{-1}^{0} x\,(1 + x)\, dx + \int_{0}^{1} x\,(1 - x)\, dx \\
&= \int_{-1}^{0} (x + x^2)\, dx + \int_{0}^{1} (x - x^2)\, dx \\
&= \frac{1}{3} - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} \\
&= 0.
\end{aligned}$$

The second moment $E(X^2)$ of X is given by

$$\begin{aligned}
E(X^2) &= \int_{-\infty}^{\infty} x^2\, f(x)\, dx \\
&= \int_{-1}^{1} x^2\,(1-|x|)\,dx \\
&= \int_{-1}^{0} x^2\,(1+x)\,dx + \int_{0}^{1} x^2\,(1-x)\,dx \\
&= \int_{-1}^{0} (x^2+x^3)\,dx + \int_{0}^{1} (x^2-x^3)\,dx \\
&= \frac{1}{3} - \frac{1}{4} + \frac{1}{3} - \frac{1}{4} \\
&= \frac{1}{6}.
\end{aligned}$$

Thus, the variance of X is given by

$$Var(X) = E(X^2) - \mu_X^2 = \frac{1}{6} - 0 = \frac{1}{6}.$$

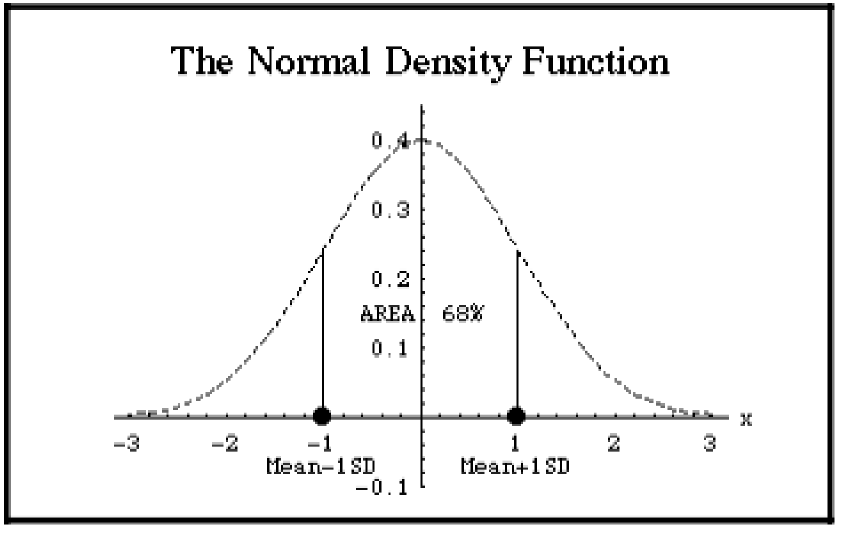
Chebyshev Inequality

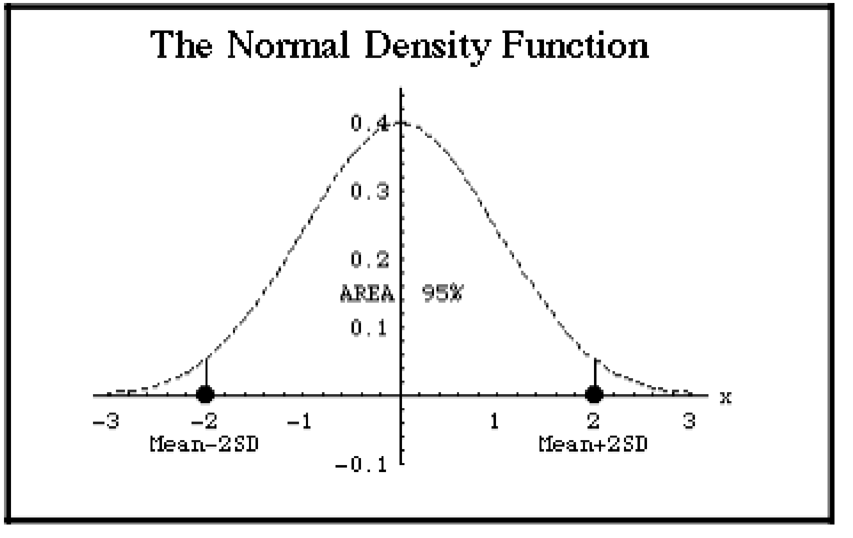
Theorem:

Let X be a random variable with probability density function f(x). If μ and σ > 0 are the mean and standard deviation of X, then

$$P(|X-\mu| < k\,\sigma) \geq 1 - \frac{1}{k^2}$$

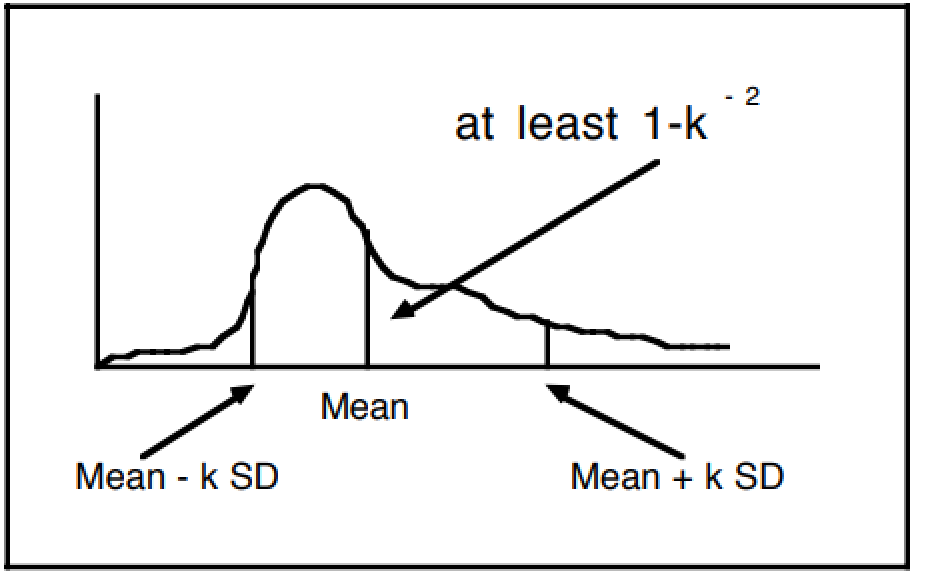
for any nonzero real positive constant k.





let mean $\mu = 0$ and the standard deviation $\sigma = 1$, and then the area between the values $\mu - \sigma$ and $\mu + \sigma$ is 68%.
Similarly, the area between the values $\mu - 2\sigma$ and $\mu + 2\sigma$ is 95%.
In this way, the standard deviation controls the area between the values $\mu - k\sigma$ and $\mu + k\sigma$ for some k if the distribution is standard normal. If we do not know the probability density function of a random variable, can we find an estimate of the area between the values $\mu - k\sigma$ and $\mu + k\sigma$ for some given k? This problem was solved by Chebychev, a well known Russian mathematician. He proved that the area under f(x) on the interval $[\mu - k\sigma, \mu + k\sigma]$ is at least $1 - k^{-2}$. This is equivalent to saying the probability that a random variable is within k standard deviations of the mean is at least $1 - k^{-2}$.

Proof: We assume that the random variable X is continuous. If X is not continuous we replace the integral by summation in the following proof. From the definition of variance, we have the following:



$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx$$

$$= \int_{-\infty}^{\mu - k\,\sigma} (x-\mu)^2 f(x)\, dx + \int_{\mu - k\,\sigma}^{\mu + k\,\sigma} (x-\mu)^2 f(x)\, dx$$

$$+ \int_{\mu + k\,\sigma}^{\infty} (x-\mu)^2 f(x)\, dx.$$

Since, $\int_{\mu-k\,\sigma}^{\mu+k\,\sigma}(x-\mu)^2\,f(x)\,dx$ is positive, we get from the above

$$\sigma^2 \geq \int_{-\infty}^{\mu-k\,\sigma}(x-\mu)^2\,f(x)\,dx + \int_{\mu+k\,\sigma}^{\infty}(x-\mu)^2\,f(x)\,dx. \tag{4.1}$$

If $x \in (-\infty,\,\mu-k\,\sigma)$, then

$$x \leq \mu - k\,\sigma.$$

Hence

$$k\,\sigma \leq \mu - x$$

for

$$k^2\,\sigma^2 \leq (\mu-x)^2.$$

That is $(\mu-x)^2 \geq k^2\,\sigma^2$. Similarly, if $x \in (\mu+k\,\sigma,\,\infty)$, then

$$x \geq \mu + k\,\sigma$$

Therefore

$$k^2\,\sigma^2 \leq (\mu - x)^2.$$

Thus if $x \notin (\mu - k\,\sigma,\, \mu + k\,\sigma)$, then

$$(\mu - x)^2 \geq k^2\,\sigma^2. \tag{4.2}$$

Using (4.2) and (4.1), we get

$$\sigma^2 \geq k^2 \sigma^2 \left[\int_{-\infty}^{\mu - k\,\sigma} f(x)\,dx + \int_{\mu + k\,\sigma}^{\infty} f(x)\,dx \right].$$

Hence

$$\frac{1}{k^2} \geq \left[\int_{-\infty}^{\mu - k\,\sigma} f(x)\,dx + \int_{\mu + k\,\sigma}^{\infty} f(x)\,dx \right].$$

Therefore

$$\frac{1}{k^2} \geq P(X \leq \mu - k\,\sigma) + P(X \geq \mu + k\,\sigma).$$

Thus

$$\frac{1}{k^2} \geq P(|X - \mu| \geq k\,\sigma)$$

which is

$$P(|X - \mu| < k\,\sigma) \geq 1 - \frac{1}{k^2}.$$

This completes the proof of this theorem.

See you next Lecture

LECTURE 13#

Example. Let the probability density function of a random variable X be

$$f(x) = \begin{cases} 630\, x^4\, (1-x)^4 & \text{if } 0 < x < 1 \\ \\ 0 & \text{otherwise.} \end{cases}$$

What is the exact value of $P(|X - \mu| \leq 2\,\sigma)$? What is the approximate value of $P(|X - \mu| \leq 2\,\sigma)$ when one uses the Chebychev inequality?

Answer: First, we find the mean and variance of the above distribution. The mean of X is given by

$$\begin{aligned} E(X) &= \int_0^1 x\, f(x)\, dx \\ &= \int_0^1 630\, x^5\, (1-x)^4\, dx \quad = 630\, \frac{5!\, 4!}{(5+4+1)!} \\ &\qquad\qquad\qquad\qquad\quad\;\; = 630\, \frac{5!\, 4!}{10!} \\ &\qquad\qquad\qquad\qquad\quad\;\; = \frac{1}{2}. \end{aligned}$$

Similarly, the variance of X can be computed from

$$
\begin{aligned}
Var(X) &= \int_0^1 x^2\, f(x)\, dx - \mu_X^2 \\
&= \int_0^1 630\, x^6\, (1-x)^4\, dx - \frac{1}{4} \quad = 630\, \frac{6!\, 4!}{(6+4+1)!} - \frac{1}{4} \quad = \frac{1}{44}.
\end{aligned}
$$

Therefore, the standard deviation of X is $\sigma = \sqrt{\frac{1}{44}} = 0.15.$

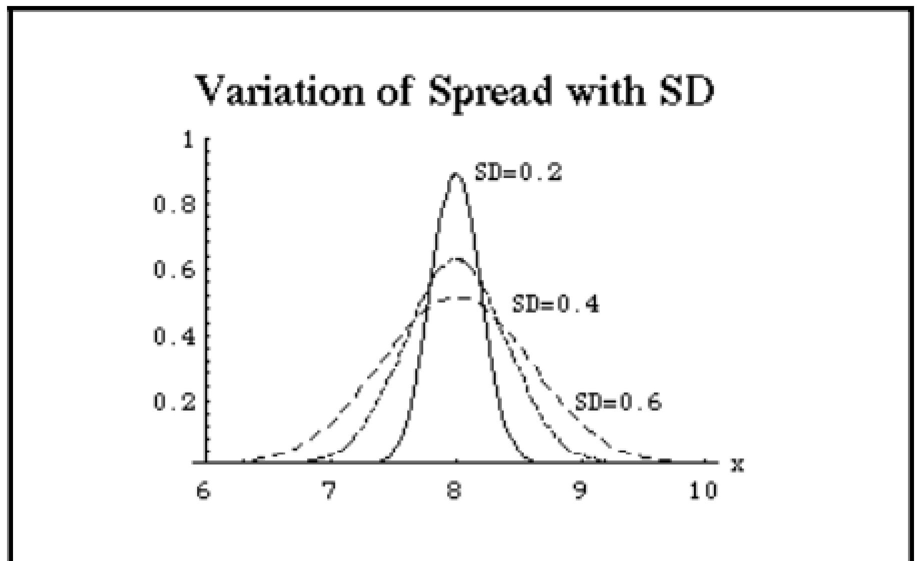
Thus

$$
\begin{aligned}
P(|X-\mu| \le 2\,\sigma) &= P(|X-0.5| \le 0.3) \\
&= P(-0.3 \le X - 0.5 \le 0.3) \\
&= P(0.2 \le X \le 0.8) \\
&= \int_{0.2}^{0.8} 630\, x^4\, (1-x)^4\, dx \\
&= 0.96.
\end{aligned}
$$

If we use the Chebyshev inequality, then we get an approximation of the exact value we have. This approximate value is

$$P(|X-\mu| \leq 2\,\sigma) \geq 1 - \frac{1}{4} = 0.75$$

Hence, Chebychev inequality tells us that if we do not know the distribution of X, then $P(|X-\mu| \leq 2\,\sigma)$ is at least 0.75.



Moment Generating Functions

Some cases , the moments are difficult to compute from the definition. A moment generating function is a real valued function from which one can generate all the moments of a given random variable. In many cases, it is easier to compute various moments of X using the moment generating function.

Definition . Let X be a random variable whose probability density function is f(x). A real valued function M :R →R defined by

$$M(t) = E\left(e^{t\,X}\right)$$

is called the moment generating function of X if this expected value exists for all t in the interval $-h < t < h$ for some $h > 0$

In general, not every random variable has a moment generating function. But if the moment generating function of a random variable exists, then it is unique.

Using the definition of expected value of a random variable, we obtain the explicit representation for M(t) as

$$M(t) = \begin{cases} \sum\limits_{x \in R_X} e^{t\,x}\, f(x) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} e^{t\,x}\, f(x)\, dx & \text{if } X \text{ is continuous.} \end{cases}$$

Example. Let X be a random variable whose moment generating function is M(t) and n be any natural number. What is the nth derivative of M(t) at t = 0?

Answer:

$$\begin{aligned} \frac{d}{dt} M(t) &= \frac{d}{dt} E\left(e^{t\,X}\right) \\ &= E\left(\frac{d}{dt} e^{t\,X}\right) \\ &= E\left(X\, e^{t\,X}\right). \end{aligned}$$

Similarly,

$$\begin{aligned} \frac{d^2}{dt^2} M(t) &= \frac{d^2}{dt^2} E\left(e^{t\,X}\right) \\ &= E\left(\frac{d^2}{dt^2} e^{t\,X}\right) \\ &= E\left(X^2\, e^{t\,X}\right). \end{aligned}$$

Hence, in general we get

$$\begin{aligned}\frac{d^n}{dt^n} M(t) &= \frac{d^n}{dt^n} E\left(e^{t\,X}\right) \\ &= E\left(\frac{d^n}{dt^n} e^{t\,X}\right) \\ &= E\left(X^n\, e^{t\,X}\right).\end{aligned}$$

If we set $t = 0$ in the n^{th} derivative, we get

$$\left.\frac{d^n}{dt^n} M(t)\right|_{t=0} = E\left(X^n\, e^{t\,X}\right)\Big|_{t=0} = E\left(X^n\right).$$

Hence the n^{th} derivative of the moment generating function of X evaluated at $t = 0$ is the n^{th} moment of X about the origin.

This example tells us if we know the moment generating function of a random variable; then we can generate all the moments of X by taking derivatives of the moment generating function and then evaluating them at zero.

Example. What is the moment generating function of the random variable X whose probability density function is given by

$$f(x) = \begin{cases} e^{-x} & \text{for } x > 0 \\ 0 & \text{otherwise?} \end{cases}$$

What are the mean and variance of X?

Answer: The moment generating function of X is

$$\begin{aligned} M(t) &= E\left(e^{t\,X}\right) \\ &= \int_0^\infty e^{t\,x}\, f(x)\, dx \quad = \int_0^\infty e^{t\,x}\, e^{-x}\, dx \\ &= \int_0^\infty e^{-(1-t)\,x}\, dx \quad = \frac{1}{1-t}\left[-e^{-(1-t)\,x}\right]_0^\infty \\ &= \frac{1}{1-t} \qquad \text{if } 1-t > 0. \end{aligned}$$

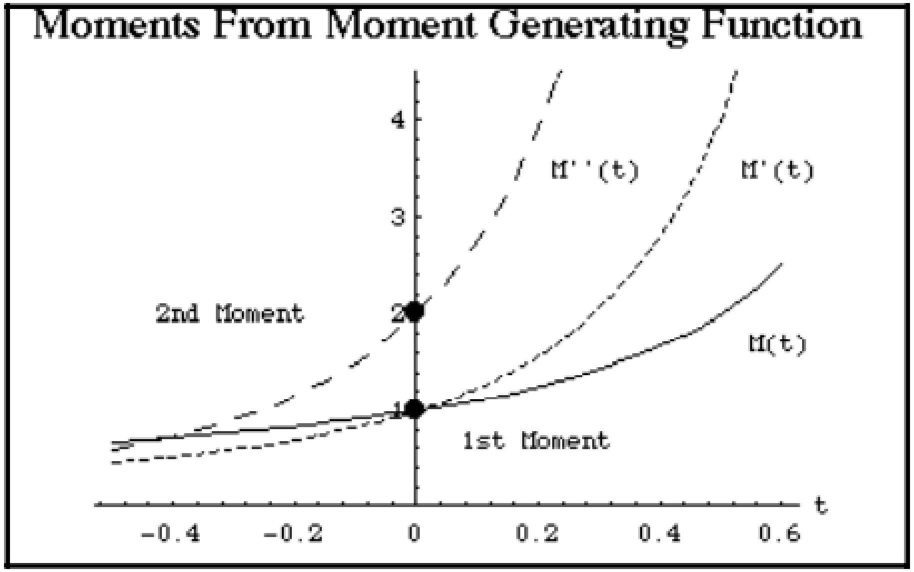
The expected value of X can be computed from M(t) as

$$
\begin{aligned}
E(X) &= \left.\frac{d}{dt}M(t)\right|_{t=0} \\
&= \left.\frac{d}{dt}(1-t)^{-1}\right|_{t=0} \quad = \left.(1-t)^{-2}\right|_{t=0} \\
&\qquad = \left.\frac{1}{(1-t)^2}\right|_{t=0} \\
&\qquad = 1.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
E(X^2) &= \left.\frac{d^2}{dt^2}M(t)\right|_{t=0} \\
&= \left.\frac{d^2}{dt^2}(1-t)^{-1}\right|_{t=0} \\
&= \left.2\,(1-t)^{-3}\right|_{t=0} \quad = \left.\frac{2}{(1-t)^3}\right|_{t=0} \\
&\qquad = 2.
\end{aligned}
$$

Therefore, the variance of X is: $Var(X) = E(X^2) - (\mu)^2 = 2 - 1 = 1.$



H.W . Let X have the probability density function

$$f(x) = \begin{cases} \frac{1}{9}\left(\frac{8}{9}\right)^x & \text{for } x = 0, 1, 2, ..., \infty \\ 0 & \text{otherwise.} \end{cases}$$

What is the moment generating function of the random variable X?

H.W. Let X be a continuous random variable with density function

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

where b > 0. If M(t) is the moment generating function of X, then what is M(−6 b)?

Theorem. Let M(t) be the moment generating function of the random variable X. If

$$M(t) = a_0 + a_1\,t + a_2\,t^2 + \cdots + a_n\,t^n + \cdots \qquad (1)$$

is the Taylor series expansion of M(t), then

$$E\left(X^n\right) = (n!)\; a_n$$

for all natural number n

Proof: Let M(t) be the moment generating function of the random variable X. The Taylor series expansion of M(t) about 0 is given by

$$M(t) = M(0) + \frac{M'(0)}{1!}t + \frac{M''(0)}{2!}t^2 + \frac{M'''(0)}{3!}t^3 + \cdots + \frac{M^{(n)}(0)}{n!}t^n + \cdots$$

Since $E(X^n) = M^{(n)}(0)$ for $n \geq 1$ and $M(0) = 1$, we have

$$M(t) = 1 + \frac{E(X)}{1!}t + \frac{E(X^2)}{2!}t^2 + \frac{E(X^3)}{3!}t^3 + \cdots + \frac{E(X^n)}{n!}t^n + \cdots \qquad (2)$$

From (1) and (2), equating the coefficients of the like powers of t, we obtain

$$a_n = \frac{E(X^n)}{n!}$$

which is

$$E(X^n) = (n!)\, a_n.$$

This proves the theorem.

Example. What is the 479th moment of X about the origin, if the moment generating function of X is $\frac{1}{1+t}$?

Answer The Taylor series expansion of M(t) = 1 /1+t can be obtained by using long division

$$
\begin{aligned}
M(t) &= \frac{1}{1+t} \\
&= \frac{1}{1-(-t)} \\
&= 1+(-t)+(-t)^2+(-t)^3+\cdots+(-t)^n+\cdots \\
&= 1-t+t^2-t^3+t^4+\cdots+(-1)^n t^n+\cdots
\end{aligned}
$$

Therefore $a_n = (-1)^n$ and from this we obtain $a_{479} = -1$.

And by above theorem, we get:

$$E\left(X^{479}\right) = (479!)\, a_{479} = -\,479!$$

Example. If the moment generating of a random variable X is $M(t) = \sum_{j=0}^{\infty} \frac{e^{(t\,j-1)}}{j!}$,

then what is the probability of the event X = 2?

Answer: By examining the given moment generating function of X, it is easy to note that X is a discrete random variable with space RX = {0, 1, 2, · · · ,∞}. Hence by definition, the moment generating function of X is

$$M(t) = \sum_{j=0}^{\infty} e^{t\,j} f(j).$$

But we are given that

$$M(t) = \sum_{j=0}^{\infty} \frac{e^{(t\,j-1)}}{j!} = \sum_{j=0}^{\infty} \frac{e^{-1}}{j!} e^{t\,j}.$$

Hence, $f(j) = \frac{e^{-1}}{j!}$ for $j = 0, 1, 2, ..., \infty.$

Thus, the probability of the event X = 2 is given by $P(X = 2) = f(2) = \frac{e^{-1}}{2!} = \frac{1}{2\,e}.$

H.W.. Let X be a random variable with

$$E\left(X^n\right) = 0.8 \qquad \text{for} \quad n = 1, 2, 3, ..., \infty.$$

What are the moment generating function and probability density function of X?

Theorem. Let X be a random variable with the moment generating function M_X (t). If a and b are any two real constants, then

$$M_{X+a}(t) = e^{a\,t}\, M_X(t) \qquad (1)$$

$$M_{b\,X}(t) = M_X(b\,t) \qquad (2)$$

$$M_{\frac{X+a}{b}}(t) = e^{\frac{a}{b}\,t}\, M_X\left(\frac{t}{b}\right). \qquad (3)$$

Proof:

$$\begin{aligned} M_{X+a}(t) &= E\left(e^{t\,(X+a)}\right) \\ &= E\left(e^{t\,X+t\,a}\right) \\ &= E\left(e^{t\,X}\,e^{t\,a}\right) \\ &= e^{t\,a}\,E\left(e^{t\,X}\right) \\ &= e^{t\,a}\,M_X(t). \end{aligned}$$

Similarly, we prove

$$\begin{aligned} M_{b\,X}(t) &= E\left(e^{t\,(b\,X)}\right) \\ &= E\left(e^{(t\,b)\,X}\right) \\ &= M_X(t\,b). \end{aligned}$$

By above cases, we easily get

$$\begin{aligned} M_{\frac{X+a}{b}}(t) &= M_{\frac{X}{b}+\frac{a}{b}}(t) \\ &= e^{\frac{a}{b}\,t}\,M_{\frac{X}{b}}(t) \\ &= e^{\frac{a}{b}\,t}\,M_X\left(\frac{t}{b}\right). \end{aligned}$$

See you next Lecture

LECTURE 14#

Example. Let the probability density function of a random variable X be

$$f(x) = \begin{cases} 630\, x^4\, (1-x)^4 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

What is the exact value of $P(|X-\mu| \leq 2\,\sigma)$? What is the approximate value of $P(|X-\mu| \leq 2\,\sigma)$ when one uses the Chebychev inequality?

Answer: First, we find the mean and variance of the above distribution. The mean of X is given by

$$\begin{aligned} E(X) &= \int_0^1 x\, f(x)\, dx \\ &= \int_0^1 630\, x^5\, (1-x)^4\, dx \quad = 630\, \frac{5!\, 4!}{(5+4+1)!} \\ &\qquad\qquad = 630\, \frac{5!\, 4!}{10!} \\ &\qquad\qquad = \frac{1}{2}. \end{aligned}$$

Similarly, the variance of X can be computed from

$$Var(X) = \int_0^1 x^2 \, f(x) \, dx - \mu_X^2$$

$$= \int_0^1 630 \, x^6 \, (1-x)^4 \, dx - \frac{1}{4} \quad = 630 \, \frac{6! \, 4!}{(6+4+1)!} - \frac{1}{4} \quad = \frac{1}{44}.$$

Therefore, the standard deviation of X is $\sigma = \sqrt{\dfrac{1}{44}} = 0.15.$

Thus

$$P(|X-\mu| \leq 2\,\sigma) = P(|X-0.5| \leq 0.3)$$

$$= P(-0.3 \leq X - 0.5 \leq 0.3)$$

$$= P(0.2 \leq X \leq 0.8)$$

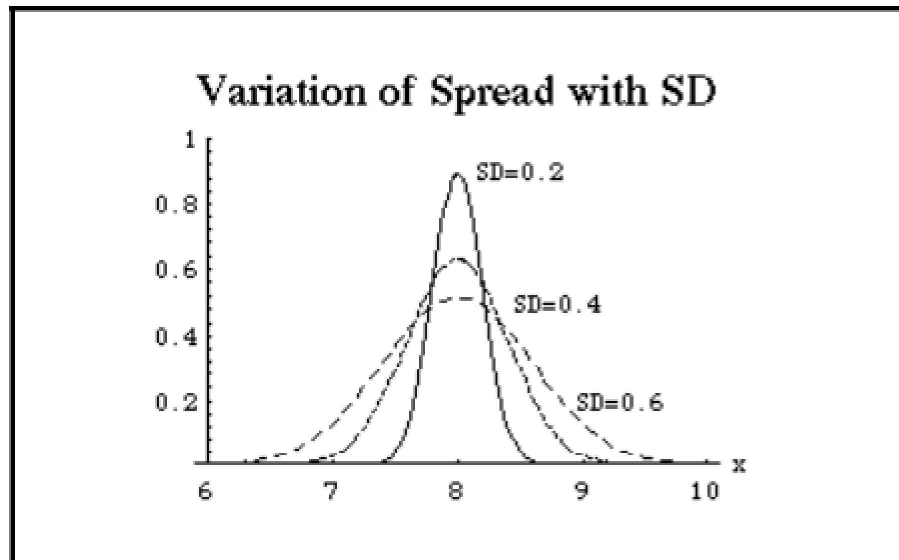
$$= \int_{0.2}^{0.8} 630 \, x^4 \, (1-x)^4 \, dx$$

$$= 0.96.$$

If we use the Chebyshev inequality, then we get an approximation of the exact value we have. This approximate value is

$$P(|X - \mu| \leq 2\,\sigma) \geq 1 - \frac{1}{4} = 0.75$$

Hence, Chebychev inequality tells us that if we do not know the distribution of X, then $P(|X - \mu| \leq 2\,\sigma)$ is at least 0.75.



Moment Generating Functions

Some cases , the moments are difficult to compute from the definition. A moment generating function is a real valued function from which one can generate all the moments of a given random variable. In many cases, it is easier to compute various moments of X using the moment generating function.

Definition . Let X be a random variable whose probability density function is f(x). A real valued function M :R →R defined by

$$M(t) = E\left(e^{t\,X}\right)$$

is called the moment generating function of X if this expected value exists for all t in the interval $-h < t < h$ for some $h > 0$

In general, not every random variable has a moment generating function. But if the moment generating function of a random variable exists, then it is unique.

Using the definition of expected value of a random variable, we obtain the explicit representation for M(t) as

$$M(t) = \begin{cases} \sum_{x \in R_X} e^{t\,x} f(x) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} e^{t\,x} f(x)\,dx & \text{if } X \text{ is continuous.} \end{cases}$$

Example. Let X be a random variable whose moment generating function is M(t) and n be any natural number. What is the nth derivative of M(t) at t = 0?

Answer:

$$\begin{aligned} \frac{d}{dt} M(t) &= \frac{d}{dt} E\left(e^{t\,X}\right) \\ &= E\left(\frac{d}{dt} e^{t\,X}\right) \\ &= E\left(X\, e^{t\,X}\right). \end{aligned}$$

Similarly,

$$\begin{aligned} \frac{d^2}{dt^2} M(t) &= \frac{d^2}{dt^2} E\left(e^{t\,X}\right) \\ &= E\left(\frac{d^2}{dt^2} e^{t\,X}\right) \\ &= E\left(X^2\, e^{t\,X}\right). \end{aligned}$$

Hence, in general we get

$$\begin{aligned}\frac{d^n}{dt^n} M(t) &= \frac{d^n}{dt^n} E\left(e^{t\,X}\right) \\ &= E\left(\frac{d^n}{dt^n} e^{t\,X}\right) \\ &= E\left(X^n\, e^{t\,X}\right).\end{aligned}$$

If we set $t = 0$ in the n^{th} derivative, we get

$$\left.\frac{d^n}{dt^n} M(t)\right|_{t=0} = E\left(X^n\, e^{t\,X}\right)\big|_{t=0} = E\left(X^n\right).$$

Hence the n^{th} derivative of the moment generating function of X evaluated at $t = 0$ is the n^{th} moment of X about the origin.

This example tells us if we know the moment generating function of a random variable; then we can generate all the moments of X by taking derivatives of the moment generating function and then evaluating them at zero.

Example. What is the moment generating function of the random variable X whose probability density function is given by

$$f(x) = \begin{cases} e^{-x} & \text{for } x > 0 \\ \\ 0 & \text{otherwise?} \end{cases}$$

What are the mean and variance of X?

Answer: The moment generating function of X is

$$\begin{aligned} M(t) &= E\left(e^{t\,X}\right) \\ &= \int_0^\infty e^{t\,x}\, f(x)\, dx \quad = \int_0^\infty e^{t\,x}\, e^{-x}\, dx \\ &\qquad\qquad\qquad\qquad = \int_0^\infty e^{-(1-t)\,x}\, dx \quad = \frac{1}{1-t}\left[-e^{-(1-t)\,x}\right]_0^\infty \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad = \frac{1}{1-t} \qquad \text{if } 1-t > 0. \end{aligned}$$

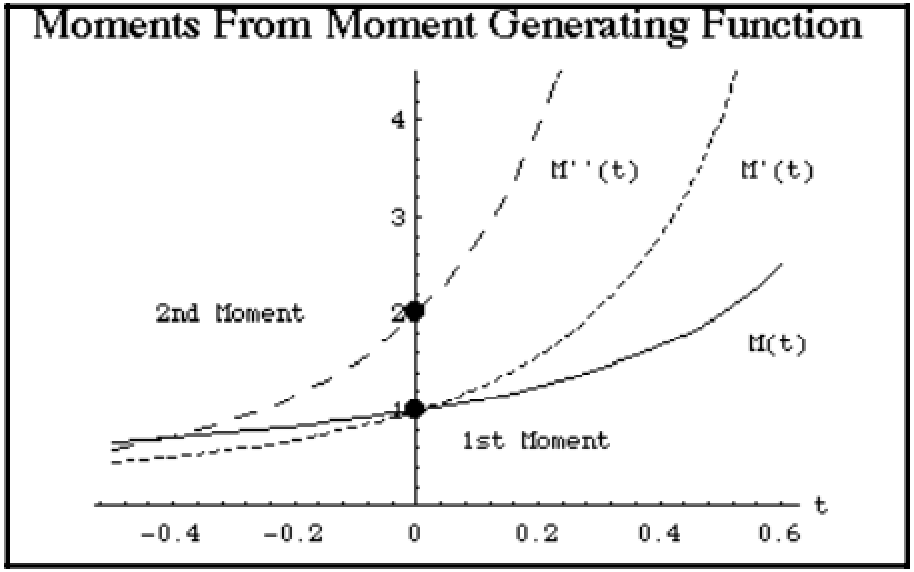
The expected value of X can be computed from M(t) as

$$
\begin{aligned}
E(X) &= \frac{d}{dt} M(t)\bigg|_{t=0} \\
&= \frac{d}{dt}(1-t)^{-1}\bigg|_{t=0} \qquad = (1-t)^{-2}\big|_{t=0} \\
&\qquad\qquad\qquad\qquad\qquad = \frac{1}{(1-t)^2}\bigg|_{t=0} \\
&\qquad\qquad\qquad\qquad\qquad = 1.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
E(X^2) &= \frac{d^2}{dt^2} M(t)\bigg|_{t=0} \\
&= \frac{d^2}{dt^2}(1-t)^{-1}\bigg|_{t=0} \\
&= 2\,(1-t)^{-3}\big|_{t=0} \qquad = \frac{2}{(1-t)^3}\bigg|_{t=0} \\
&\qquad\qquad\qquad\qquad\qquad = 2.
\end{aligned}
$$

Therefore, the variance of X is: $Var(X) = E(X^2) - (\mu)^2 = 2 - 1 = 1.$



H.W . Let X have the probability density function

$$f(x) = \begin{cases} \frac{1}{9} \left(\frac{8}{9}\right)^x & \text{for } x = 0, 1, 2, ..., \infty \\ 0 & \text{otherwise.} \end{cases}$$

What is the moment generating function of the random variable X?

H.W. Let X be a continuous random variable with density function

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

where b > 0. If M(t) is the moment generating function of X, then what is M(−6 b)?

Theorem. Let M(t) be the moment generating function of the random variable X. If

$$M(t) = a_0 + a_1\, t + a_2\, t^2 + \cdots + a_n\, t^n + \cdots \qquad (1)$$

is the Taylor series expansion of M(t), then

$$E\left(X^n\right) = (n!)\ a_n$$

for all natural number n

Proof: Let M(t) be the moment generating function of the random variable X. The Taylor series expansion of M(t) about 0 is given by

$$M(t) = M(0) + \frac{M'(0)}{1!}\,t + \frac{M''(0)}{2!}\,t^2 + \frac{M'''(0)}{3!}\,t^3 + \cdots + \frac{M^{(n)}(0)}{n!}\,t^n + \cdots$$

Since $E(X^n) = M^{(n)}(0)$ for $n \geq 1$ and $M(0) = 1$, we have

$$M(t) = 1 + \frac{E(X)}{1!}\,t + \frac{E(X^2)}{2!}\,t^2 + \frac{E(X^3)}{3!}\,t^3 + \cdots + \frac{E(X^n)}{n!}\,t^n + \cdots \qquad (2)$$

From (1) and (2), equating the coefficients of the like powers of t, we obtain

$$a_n = \frac{E\left(X^n\right)}{n!}$$

which is

$$E\left(X^n\right) = (n!)\; a_n.$$

This proves the theorem.

Example. What is the 479th moment of X about the origin, if the moment generating function of X is $\frac{1}{1+t}$?

Answer The Taylor series expansion of M(t) = 1 /1+t can be obtained by using long division

$$\begin{aligned} M(t) &= \frac{1}{1+t} \\ &= \frac{1}{1-(-t)} \\ &= 1 + (-t) + (-t)^2 + (-t)^3 + \cdots + (-t)^n + \cdots \\ &= 1 - t + t^2 - t^3 + t^4 + \cdots + (-1)^n t^n + \cdots \end{aligned}$$

Therefore $a_n = (-1)^n$ and from this we obtain $a_{479} = -1$.

And by above theorem, we get:

$$E\left(X^{479}\right) = (479!)\, a_{479} = -\,479!$$

Example. If the moment generating of a random variable X is $M(t)=\sum_{j=0}^{\infty}\frac{e^{(t\,j-1)}}{j!}$,

then what is the probability of the event X = 2?

Answer: By examining the given moment generating function of X, it is easy to note that X is a discrete random variable with space RX = {0, 1, 2, · · · ,∞}. Hence by definition, the moment generating function of X is

$$M(t)=\sum_{j=0}^{\infty} e^{t\,j}\, f(j).$$

But we are given that

$$M(t)=\sum_{j=0}^{\infty}\frac{e^{(t\,j-1)}}{j!} = \sum_{j=0}^{\infty}\frac{e^{-1}}{j!}\, e^{t\,j}.$$

Hence, $f(j)=\frac{e^{-1}}{j!}$ for $j=0,1,2,\ldots,\infty.$

Thus, the probability of the event X = 2 is given by $P(X=2)=f(2)=\frac{e^{-1}}{2!}=\frac{1}{2\,e}.$

H.W.. Let X be a random variable with

$$E\left(X^{n}\right)=0.8 \qquad \text{for} \quad n=1,2,3,\ldots,\infty.$$

What are the moment generating function and probability density function of X?

Theorem. Let X be a random variable with the moment generating function M_X (t). If a and b are any two real constants, then

$$M_{X+a}(t)=e^{a\,t}\,M_X(t) \qquad (1)$$

$$M_{b\,X}(t)=M_X(b\,t) \qquad (2)$$

$$M_{\frac{X+a}{b}}(t)=e^{\frac{a}{b}\,t}\,M_X\left(\frac{t}{b}\right). \qquad (3)$$

Proof:
$$\begin{aligned} M_{X+a}(t) &= E\left(e^{t\,(X+a)}\right) \\ &= E\left(e^{t\,X+t\,a}\right) \\ &= E\left(e^{t\,X}\,e^{t\,a}\right) \\ &= e^{t\,a}\,E\left(e^{t\,X}\right) \\ &= e^{t\,a}\,M_X(t). \end{aligned}$$

Similarly, we prove
$$\begin{aligned} M_{b\,X}(t) &= E\left(e^{t\,(b\,X)}\right) \\ &= E\left(e^{(t\,b)\,X}\right) \\ &= M_X(t\,b). \end{aligned}$$

By above cases, we easily get
$$\begin{aligned} M_{\frac{X+a}{b}}(t) &= M_{\frac{X}{b}+\frac{a}{b}}(t) \\ &= e^{\frac{a}{b}\,t}\,M_{\frac{X}{b}}(t) \\ &= e^{\frac{a}{b}\,t}\,M_X\left(\frac{t}{b}\right). \end{aligned}$$

The distribution function for a random variable X is

$$F(x) = \begin{cases} 1 - e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Find (a) the density function, (b) the probability that $X > 2$, and (c) the probability that $-3 < X \leq 4$.

A continuous random variable X has probability density given by

$$f(x) = \begin{cases} 2e^{-2x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Find (a) $E(X)$, (b) $E(X^2)$.

See you next Lecture

LECTURE 15#

Joint Distributions

2. CONTINUOUS CASE. The case where both variables are continuous is obtained easily by analogy with the discrete case on replacing sums by integrals. Thus the joint probability function for the random variables X and Y (or, as it is more commonly called, the joint density function of X and Y) is defined by

1. $f(x, y) \geq 0$
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$

Example. Let the joint density function of X and Y be given by

$$f(x,y) = \begin{cases} k\,xy^2 & \text{if } 0 < x < y < 1 \\ 0 & \text{otherwise.} \end{cases}$$

What is the value of the constant k ?

Answer: Since f is a joint probability density function, we have

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

$$= \int_0^1 \int_0^y k\, x\, y^2\, dx\, dy$$

$$= \int_0^1 k\, y^2 \int_0^y x\, dx\, dy$$

$$= \frac{k}{2} \int_0^1 y^4\, dy$$

$$= \frac{k}{10} \left[y^5 \right]_0^1$$

$$= \frac{k}{10}.$$

Hence $k = 10$.

Note :If we know the joint probability density function f of the random variables X and Y , then we can compute the probability of the event A from

$$P(A) = \int\int_A f(x, y)\, dx\, dy.$$

Example :Let the joint density of the continuous random variables X and Y be

$$f(x, y) = \begin{cases} \frac{6}{5}\,(x^2 + 2\,xy) & \text{if } 0 \le x \le 1;\ 0 \le y \le 1 \\ 0 & \text{elsewhere.} \end{cases}$$

What is the probability of the event $(X \le Y)$?

Answer: Let $A = (X \leq Y)$. we want to find

$$
\begin{aligned}
P(A) &= \int\int_A f(x,\, y)\, dx\, dy \\
&= \int_0^1 \left[\int_0^y \frac{6}{5}\, \left(x^2 + 2\, x\, y\right)\, dx\right]\, dy \\
&= \frac{6}{5} \int_0^1 \left[\frac{x^3}{3} + x^2\, y\right]_{x=0}^{x=y} dy \\
&= \frac{6}{5} \int_0^1 \frac{4}{3}\, y^3\, dy \\
&= \frac{2}{5}\, \left[y^4\right]_0^1 \\
&= \frac{2}{5}.
\end{aligned}
$$

Definition : Let (X, Y) be a continuous bivariate random variable. Let f(x, y) be the joint probability density function of X and Y . The function

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

is called the **marginal probability density function of X**. Similarly, the function

$$f_2(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

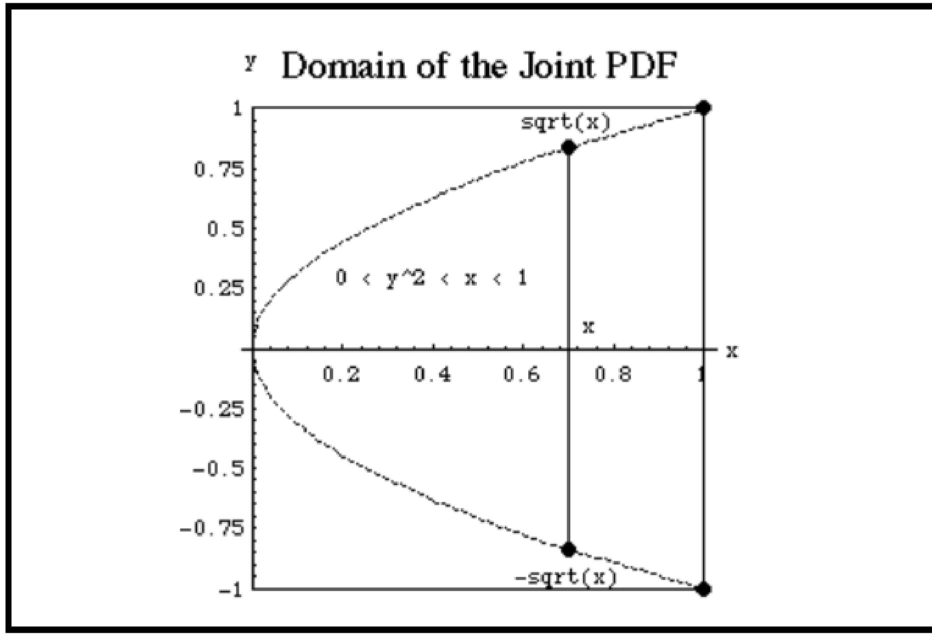
is called the **marginal probability density function of Y** .

Example : If the joint density function for X and Y is given by

$$f(x, y) = \begin{cases} \frac{3}{4} & \text{for } 0 < y^2 < x < 1 \\ 0 & \text{otherwise,} \end{cases}$$

then what is the marginal density function of X, for 0 < x < 1?

Answer: The domain of the f consists of the region bounded by the curve $x = y^2$ and the vertical line x = 1



Hence

$$f_1(x) = \int_{-\sqrt{x}}^{\sqrt{x}} \frac{3}{4}\, dy$$

$$= \left[\frac{3}{4}\, y\right]_{-\sqrt{x}}^{\sqrt{x}} \quad = \frac{3}{2}\sqrt{x}.$$

Example : Let X and Y have joint density function

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

What is the marginal density of X where nonzero?

Answer: The marginal density of X is given by

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \int_{x}^{\infty} 2\, e^{-x-y}\, dy$$

$$= 2\, e^{-x} \int_{x}^{\infty} e^{-y}\, dy \; = 2\, e^{-x} \left[-e^{-y} \right]_{x}^{\infty} = 2\, e^{-x}\, e^{-x}$$

$$= 2\, e^{-2x} \qquad 0 < x < \infty.$$

Definition: Let X and Y be the continuous random variables with joint probability density function f(x, y). The joint cumulative distribution function F(x, y) of X and Y is defined as

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

From the fundamental theorem of calculus, we again obtain $f(x,\, y) = \dfrac{\partial^2 F}{\partial x\, \partial y}.$

Example:. If the joint cumulative distribution function of X and Y is given by

$$F(x, y) = \begin{cases} \frac{1}{5}\,\left(2\,x^3\,y + 3\,x^2\,y^2\right) & \text{for } 0 < x, y < 1 \\ 0 & \text{elsewhere,} \end{cases}$$

then what is the joint density of X and Y ?

Answer:

$$
\begin{aligned}
f(x,\, y) &= \frac{1}{5}\,\frac{\partial}{\partial x}\,\frac{\partial}{\partial y}\left(2\,x^3\,y + 3\,x^2\,y^2\right) \\
&= \frac{1}{5}\,\frac{\partial}{\partial x}\left(2\,x^3 + 6\,x^2\,y\right) \\
&= \frac{1}{5}\,\left(6\,x^2 + 12\,x\,y\right) \\
&= \frac{6}{5}\,(x^2 + 2\,x\,y).
\end{aligned}
$$

Hence, the joint density of X and Y is given by

$$
f(x,y) = \begin{cases} \frac{6}{5}\,\left(x^2 + 2\,x\,y\right) & \text{for } 0 < x, y < 1 \\[2em] 0 & \text{elsewhere.} \end{cases}
$$

See you next Lecture

LECTURE 16#

Conditional Distributions

Definition. Let X and Y be any two random variables with joint density f(x, y) and marginals f1(x) and f2(y). The conditional probability density function g of X, given (the event) Y = y, is defined as

$$g(x\,/\,y) = \frac{f(x,\,y)}{f_2(y)} \qquad f_2(y) > 0.$$

Similarly, the conditional probability density function h of Y , given (the event) X = x, is defined as

$$h(y\,/\,x) = \frac{f(x,\,y)}{f_1(x)} \qquad f_1(x) > 0.$$

Example. Let X and Y be discrete random variables with joint probability function

$$f(x,y) = \begin{cases} \frac{1}{21}(x+y) & \text{for } x = 1,2,3;\ y = 1,2. \\ 0 & \text{elsewhere.} \end{cases}$$

What is the conditional probability density function of X, given Y = 2 ?

Answer: We want to find $g(x/2)$. Since

$$g(x\,/\,2) = \frac{f(x,\,2)}{f_2(2)}$$

we should first compute the marginal of Y, that is $f_2(2)$. The marginal of Y is given by

$$\begin{aligned} f_2(y) &= \sum_{x=1}^{3} \frac{1}{21}(x+y) \\ &= \frac{1}{21}(6+3\,y). \end{aligned}$$

Hence $f_2(2) = \frac{12}{21}$. Thus, the conditional probability density function of X, given $Y = 2$, is

$$\begin{aligned} g(x/2) &= \frac{f(x,\, 2)}{f_2(2)} \\ &= \frac{\frac{1}{21}\,(x+2)}{\frac{12}{21}} \\ &= \frac{1}{12}\,(x+2), \qquad\qquad x = 1, 2, 3. \end{aligned}$$

Example :Let X and Y be discrete random variables with joint probability density function

$$f(x,y) = \begin{cases} \frac{x+y}{32} & \text{for } x = 1, 2;\ y = 1, 2, 3, 4 \\ \\ 0 & \text{otherwise.} \end{cases}$$

What is the conditional probability of Y given X = x ?

Answer:

$$\begin{aligned} f_1(x) &= \sum_{y=1}^{4} f(x,y) \\ &= \frac{1}{32} \sum_{y=1}^{4} (x+y) \\ &= \frac{1}{32}\,(4\,x+10). \end{aligned}$$

Therefore

$$\begin{aligned} h(y/x) &= \frac{f(x,y)}{f_1(x)} \\ &= \frac{\frac{1}{32}\,(x+y)}{\frac{1}{32}\,(4\,x+10)} \\ &= \frac{x+y}{4x+10}. \end{aligned}$$

Thus, the conditional probability Y given $X = x$ is

$$h(y/x) = \begin{cases} \frac{x+y}{4x+10} & \text{for } x = 1, 2;\; y = 1, 2, 3, 4 \\ 0 & \text{otherwise.} \end{cases}$$

H.W : Let X and Y be continuous random variables with joint pdf

$$f(x,y) = \begin{cases} 12\,x & \text{for } 0 < y < 2x < 1 \\ 0 & \text{otherwise .} \end{cases}$$

What is the conditional density function of Y given X = x ?

Independence of Random Variables

we define the concept of stochastic independence of two random variables X and Y . The conditional probability density function g of X given Y = y usually depends on y. If g is independent of y, then the random variables X and Y are said to be independent. This motivates the following definition

Definition. Let X and Y be any two random variables with joint density f(x, y) and marginals f1(x) and f2(y). The random variables X and Y are (stochastically) independent if and only if

f(x, y) = f1(x) f2(y)

Example. Let X and Y be discrete random variables with joint density

$$f(x,y) = \begin{cases} \frac{1}{36} & \text{for } 1 \leq x = y \leq 6 \\ \frac{2}{36} & \text{for } 1 \leq x < y \leq 6. \end{cases}$$

Are X and Y stochastically independent?

Answer: The marginals of X and Y are given by

$$\begin{aligned} f_1(x) &= \sum_{y=1}^{6} f(x,\, y) \\ &= f(x,x) + \sum_{y>x} f(x,y) + \sum_{y<x} f(x,y) \\ &= \frac{1}{36} + (6-x)\,\frac{2}{36} + 0 \\ &= \frac{13-2x}{36}, \qquad \text{for} \quad x = 1, 2, ..., 6 \end{aligned}$$

and

$$\begin{aligned} f_2(y) &= \sum_{x=1}^{6} f(x,\, y) \\ &= f(y,y) + \sum_{x<y} f(x,y) + \sum_{x>y} f(x,y) \\ &= \frac{1}{36} + (y-1)\,\frac{2}{36} + 0 \\ &= \frac{2y-1}{36}, \qquad \text{for} \quad y = 1, 2, ..., 6. \end{aligned}$$

Since

$$f(1,1) = \frac{1}{36} \neq \frac{11}{36}\,\frac{1}{36} = f_1(1)\, f_2(1),$$

we conclude that $f(x,y) \neq f_1(x)\, f_2(y)$, and X and Y are not independent.

However, if one knows the marginals of X and Y , then it is not possible to find the joint density of X and Y unless the random variables are independent

Example : Let X and Y have the joint density

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

Are X and Y stochastically independent?

Answer: The marginals of X and Y are given by

$$f_1(x) = \int_0^\infty f(x, y)\, dy = \int_0^\infty e^{-(x+y)}\, dy = e^{-x}$$

and

$$f_2(y) = \int_0^\infty f(x, y)\, dx = \int_0^\infty e^{-(x+y)}\, dx = e^{-y}.$$

Hence

$$f(x, y) = e^{-(x+y)} = e^{-x}\, e^{-y} = f_1(x)\, f_2(y).$$

Thus, X and Y are stochastically independent.

If X and Y are independent, then the random variables U = ϕ(X) and V = ψ(Y) are also independent. Here ϕ, ψ : RI → RI are some real valued functions

Definition :The random variables X and Y are said to be independent and identically distributed (IID) if and only if they are independent and have the same distribution.

Example. Let X and Y be two independent random variables with identical probability density function given by

$$f(x) = \begin{cases} e^{-x} & \text{for } x > 0 \\ 0 & \text{elsewhere.} \end{cases}$$

What is the probability density function of W = min{X, Y } ?

Answer: Let $G(w)$ be the cumulative distribution function of W. Then

$$\begin{aligned}
G(w) &= P(W \leq w) \\
&= 1 - P(W > w) \\
&= 1 - P(\min\{X, Y\} > w) \\
&= 1 - P(X > w \,\text{and}\, Y > w) \\
&= 1 - P(X > w)\, P(Y > w) \qquad \text{(since } X \text{ and } Y \text{ are independent)} \\
&= 1 - \left(\int_w^\infty e^{-x}\, dx \right) \left(\int_w^\infty e^{-y}\, dy \right) \\
&= 1 - \left(e^{-w} \right)^2 \\
&= 1 - e^{-2w}.
\end{aligned}$$

Thus, the probability density function of W is

$$g(w) = \frac{d}{dw} G(w) = \frac{d}{dw} \left(1 - e^{-2w} \right) = 2\, e^{-2w}.$$

Hence

$$g(w) = \begin{cases} 2\, e^{-2w} & \text{for } w > 0 \\ 0 & \text{elsewhere.} \end{cases}$$

See you next Lecture

LECTURE 17#

Covariance of Bivariate Random Variables

Definition. Let X and Y be any two random variables with joint density function f(x, y). The product moment of X and Y , denoted by E(XY), is defined as

$$E(XY) = \begin{cases} \sum_{x \in R_X} \sum_{y \in R_Y} xy\, f(x, y) & \text{if } X \text{ and } Y \text{ are discrete} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\, f(x, y)\, dx\, dy & \text{if } X \text{ and } Y \text{ are continuous.} \end{cases}$$

Here, R_X and R_Y represent the range spaces of X and Y respectively.

Definition. Let X and Y be any two random variables with joint density function f(x, y). The covariance between X and Y , denoted by Cov(X, Y) (or σ_{XY}), is defined as

$$Cov(X, Y) = E((X - \mu_X)(Y - \mu_Y)),$$

where μ_X and μ_Y are mean of X and Y, respectively.

Notice that the covariance of X and Y is really the product moment of $X - \mu_X$ and $Y - \mu_Y$. Further, the mean of μ_X is given by

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x\, f_1(x)\, dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f(x,y)\, dx\, dy,$$

and similarly the mean of Y is given by

$$\mu_Y = E(Y) = \int_{-\infty}^{\infty} y\, f_2(y)\, dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\, f(x,y)\, dy\, dx.$$

Theorem. Let X and Y be any two random variables. Then Cov(X, Y) = E(XY) − E(X) E(Y).

Proof:

$$\begin{aligned}
Cov(X, Y) &= E((X-\mu_X)(Y-\mu_Y)) \\
&= E(XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y) \\
&= E(XY) - \mu_X E(Y) - \mu_Y E(X) + \mu_X \mu_Y \\
&= E(XY) - \mu_X \mu_Y - \mu_Y \mu_X + \mu_X \mu_Y \\
&= E(XY) - \mu_X \mu_Y \\
&= E(XY) - E(X)\, E(Y).
\end{aligned}$$

Corollary : $Cov(X,X) = \sigma_X^2.$

Proof:

$$\begin{aligned} Cov(X,X) &= E(XX) - E(X)\,E(X) \\ &= E(X^2) - \mu_X^2 \\ &= Var(X) \\ &= \sigma_X^2. \end{aligned}$$

Example. Let X and Y be discrete random variables with joint density

$$f(x,y) = \begin{cases} \frac{x+2y}{18} & \text{for } x = 1,2;\ y = 1,2 \\ 0 & \text{elsewhere.} \end{cases}$$

What is the covariance σ_{XY} between X and Y.

Answer: The marginal of X is

$$f_1(x) = \sum_{y=1}^{2} \frac{x+2y}{18} = \frac{1}{18}\,(2x+6).$$

Hence the expected value of X is

$$\begin{aligned} E(X) &= \sum_{x=1}^{2} x\, f_1(x) \\ &= 1\, f_1(1) + 2 f_1(2) \\ &= \frac{8}{18} + 2\,\frac{10}{18} \\ &= \frac{28}{18}. \end{aligned}$$

Similarly, the marginal of Y is

$$f_2(y) = \sum_{x=1}^{2} \frac{x+2y}{18} = \frac{1}{18}\,(3+4y).$$

Hence the expected value of Y is

$$\begin{aligned}
E(Y) &= \sum_{y=1}^{2} y\, f_2(y) \\
&= 1\, f_2(1) + 2 f_2(2) \\
&= \frac{7}{18} + 2\, \frac{11}{18} \\
&= \frac{29}{18}.
\end{aligned}$$

Further, the product moment of X and Y is given by

$$\begin{aligned}
E(XY) &= \sum_{x=1}^{2} \sum_{y=1}^{2} x\, y\, f(x,\, y) \\
&= f(1,1) + 2\, f(1,2) + 2\, f(2,1) + 4\, f(2,2) \\
&= \frac{3}{18} + 2\, \frac{5}{18} + 2\, \frac{4}{18} + 4\, \frac{6}{18} \\
&= \frac{3 + 10 + 8 + 24}{18} \\
&= \frac{45}{18}.
\end{aligned}$$

Hence, the covariance between X and Y is given by

$$\begin{aligned} Cov(X, Y) &= E(XY) - E(X)\,E(Y) \\ &= \frac{45}{18} - \left(\frac{28}{18}\right)\left(\frac{29}{18}\right) \\ &= \frac{(45)\,(18) - (28)\,(29)}{(18)\,(18)} \\ &= \frac{810 - 812}{324} \\ &= -\frac{2}{324} = -0.00617. \end{aligned}$$

Note: The covariance between two random variables may be negative

H.W. Let X and Y have the joint density function

$$f(x,y) = \begin{cases} x + y & \text{if } 0 < x, y < 1 \\ 0 & \text{elsewhere .} \end{cases}$$

What is the covariance between X and Y ?

Theorem. If X and Y are any two random variables and a, b, c, and d are real constants, then

$$Cov(a\,X+b,\,c\,Y+d)=a\,c\,Cov(X,Y).$$

Proof:

$$
\begin{aligned}
&Cov(a\,X+b,\,c\,Y+d)\\
&\quad= E\left((aX+b)(cY+d)\right)-E(aX+b)\,E(cY+d)\\
&\quad= E\left(acXY+adX+bcY+bd\right)-\left(aE(X)+b\right)\left(cE(Y)+d\right)\\
&\quad= ac\,E(XY)+ad\,E(X)+bc\,E(Y)+bd\\
&\qquad\quad -\left[ac\,E(X)\,E(Y)+ad\,E(X)+bc\,E(Y)+bd\right]\\
&\quad= ac\,\left[E(XY)-E(X)\,E(Y)\right]\\
&\quad= ac\,Cov(X,\,Y).
\end{aligned}
$$

Example. If the product moment of X and Y is 3 and the mean of X and Y are both equal to 2, then what is the covariance of the random variables 2X + 10 and (−5 /2) Y + 3 ?

Example : Let X and Y have the joint density

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

Are X and Y stochastically independent?

Answer: Since $E(XY) = 3$ and $E(X) = 2 = E(Y)$, the covariance of X and Y is given by

$$Cov(X, Y) = E(XY) - E(X)\,E(Y) = 3 - 4 = -1.$$

Then the covariance of $2X + 10$ and $-\frac{5}{2}\,Y + 3$ is given by

$$\begin{aligned} Cov\left(2X + 10, -\frac{5}{2}Y + 3\right) &= 2\left(-\frac{5}{2}\right) Cov(X, Y) \\ &= (-5)\,(-1) \\ &= 5. \end{aligned}$$

Remark. Notice that the above Theorem can be furthered improved. That is, if X, Y , Z are three random variables, then

$$Cov(X+Y,Z) = Cov(X,Z) + Cov(Y,Z)$$

and

$$Cov(X,Y+Z) = Cov(X,Y) + Cov(X,Z).$$

The first formula can be established as follows. Consider

$$\begin{aligned}
Cov(X+Y,Z) &= E((X+Y)Z) - E(X+Y)\,E(Z) \\
&= E(XZ+YZ) - E(X)E(Z) - E(Y)E(Z) \\
&= E(XZ) - E(X)E(Z) + E(YZ) - E(Y)E(Z) \\
&= Cov(X,Z) + Cov(Y,Z).
\end{aligned}$$

Theorem. If X and Y are independent random variables, then

$$E(XY) = E(X)\,E(Y).$$

Proof: Recall that X and Y are independent if and only if

$$f(x,y) = f_1(x)\, f_2(y).$$

Let us assume that X and Y are continuous. Therefore

$$\begin{aligned} E(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, y\, f(x,y)\, dx\, dy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, y\, f_1(x)\, f_2(y)\, dx\, dy \\ &= \left(\int_{-\infty}^{\infty} x\, f_1(x)\, dx\right)\left(\int_{-\infty}^{\infty} y\, f_2(y)\, dy\right) \\ &= E(X)\, E(Y). \end{aligned}$$

If X and Y are discrete, then replace the integrals by appropriate sums to prove the same result.

Example. Let X and Y be two independent random variables with respective densitv functions:

$$f(x) = \begin{cases} 3\,x^2 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

What is $E\left(\frac{X}{Y}\right)$?

and

$$g(y) = \begin{cases} 4\,y^3 & \text{if } 0 < y < 1 \\ 0 & \text{otherwise .} \end{cases}$$

Answer: Since X and Y are independent, the joint density of X and Y is given by

$$h(x,y) = f(x)\,g(y).$$

Therefore

$$\begin{aligned}
E\left(\frac{X}{Y}\right) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{x}{y}\,h(x,y)\,dx\,dy \\
&= \int_0^1\int_0^1 \frac{x}{y}\,f(x)\,g(y)\,dx\,dy \\
&= \int_0^1\int_0^1 \frac{x}{y}\,3x^2\,4y^3\,dx\,dy \\
&= \left(\int_0^1 3x^3\,dx\right)\left(\int_0^1 4y^2\,dy\right) \\
&= \left(\frac{3}{4}\right)\,\left(\frac{4}{3}\right) = 1.
\end{aligned}$$

Remark: The independence of X and Y does not imply $E\left(\frac{X}{Y}\right) = \frac{E(X)}{E(Y)}$ but only implies $E\left(\frac{X}{Y}\right) = E(X)\,E\left(Y^{-1}\right)$. Further, note that $E\left(Y^{-1}\right)$ is not equal to $\frac{1}{E(Y)}$.

Theorem. If X and Y are independent random variables, then the covariance between X and Y is always zero, that is Cov(X, Y) = 0.

Proof: Suppose X and Y are independent, then by above theorem, we have E(XY) = E(X) E(Y). Consider

$$\begin{aligned} Cov(X, Y) &= E(XY) - E(X)\,E(Y) \\ &= E(X)\,E(Y) - E(X)\,E(Y) \\ &= 0. \end{aligned}$$

Example. Let the random variables X and Y have the joint density

$$f(x,y) = \begin{cases} \frac{1}{4} & \text{if } (x,y) \in \{\,(0,1),\ (0,-1),\ (1,0),\ (-1,0)\,\} \\ 0 & \text{otherwise.} \end{cases}$$

What is the covariance of X and Y ? Are the random variables X and Y independent?

Answer: The joint density of X and Y are shown in the following table with the marginals $f_1(x)$ and $f_2(y)$.

(x,y)	-1	0	1	$f_2(y)$
-1	0	$\frac{1}{4}$	0	$\frac{1}{4}$
0	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{2}{4}$
1	0	$\frac{1}{4}$	0	$\frac{1}{4}$
$f_1(x)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	

From this table, we see that

$$0 = f(0,0) \neq f_1(0)\, f_2(0) = \left(\frac{2}{4}\right)\, \left(\frac{2}{4}\right) = \frac{1}{4}$$

and thus

$$f(x,y) \neq f_1(x)\, f_2(y)$$

for all (x, y) is the range space of the joint variable (X, Y). Therefore X and Y are not independent. Next, we compute the covariance between X and Y . For this we need

$E(X)$, $E(Y)$ and $E(XY)$. The expected value of X is

$$E(X) = \sum_{x=-1}^{1} x f_1(x) \;= (-1)\, f_1(-1) + (0) f_1(0) + (1)\, f_1(1)$$

$$= -\frac{1}{4} + 0 + \frac{1}{4}$$

$$= 0.$$

Similarly, the expected value of Y is

$$E(Y) = \sum_{y=-1}^{1} y f_2(y) \;= (-1)\, f_2(-1) + (0) f_2(0) + (1)\, f_2(1)$$

$$= -\frac{1}{4} + 0 + \frac{1}{4} \;= 0.$$

The product moment of X and Y is given by

$$
\begin{aligned}
E(XY) &= \sum_{x=-1}^{1} \sum_{y=-1}^{1} x\, y\, f(x,y) \\
&= (1)\, f(-1,-1) + (0)\, f(-1,0) + (-1)\, f(-1,1) \\
&+ (0)\, f(0,-1) + (0)\, f(0,0) + (0)\, f(0,1) \\
&+ (-1)\, f(1,-1) + (0)\, f(1,0) + (1)\, f(1,1) \\
&= 0.
\end{aligned}
$$

Hence, the covariance between X and Y is given by

$$
Cov(X,Y) = E(XY) - E(X)\, E(Y) = 0.
$$

Remark. This example shows that if the covariance of X and Y is zero that does not mean the random variables are independent. However, we know from Theorem that if X and Y are independent, then the Cov(X, Y) is always zero.

See you next Lecture

LECTURE 18#

Variance of the Linear Combination of Random Variables

Theorem. Let X and Y be any two random variables and let a and b be any two real numbers. Then

$$Var(aX + bY) = a^2\, Var(X) + b^2\, Var(Y) + 2\,a\,b\,Cov(X, Y).$$

Proof:

$$\begin{aligned}
&Var(aX + bY)\\
&= E\left([aX + bY - E(aX + bY)]^2\right)\\
&= E\left([aX + bY - a\,E(X) - b\,E(Y)]^2\right)\\
&= E\left([a\,(X - \mu_X) + b\,(Y - \mu_Y)]^2\right)\\
&= E\left(a^2\,(X - \mu_X)^2 + b^2\,(Y - \mu_Y)^2 + 2\,a\,b\,(X - \mu_X)\,(Y - \mu_Y)\right)\\
&= a^2\,E\left((X - \mu_X)^2\right) + b^2\,E\left((X - \mu_X)^2\right) + 2\,a\,b\,E((X - \mu_X)\,(Y - \mu_Y))\\
&= a^2\,Var(X) + b^2\,Var(Y) + 2\,a\,b\,Cov(X, Y).
\end{aligned}$$

Example. If $Var(X + Y) = 3$, $Var(X - Y) = 1$, $E(X) = 1$ and $E(Y) = 2$, then what is $E(XY)$?

Answer:

$$Var(X+Y) = \sigma_X^2 + \sigma_Y^2 + 2\,Cov(X,Y),$$
$$Var(X-Y) = \sigma_X^2 + \sigma_Y^2 - 2\,Cov(X,Y).$$

Hence, we get

$$\begin{aligned} Cov(X,Y) &= \frac{1}{4}\left[Var(X+Y) - Var(X-Y)\right] \\ &= \frac{1}{4}\left[3-1\right] \\ &= \frac{1}{2}. \end{aligned}$$

Therefore, the product moment of X and Y is given by

$$\begin{aligned} E(XY) &= Cov(X,Y) + E(X)\,E(Y) \\ &= \frac{1}{2} + (1)\,(2) \\ &= \frac{5}{2}. \end{aligned}$$

Example. Let X and Y be random variables with V ar(X) = 4, V ar(Y) = 9 and V ar(X − Y) = 16. What is Cov(X, Y) ?

Answer:

$$Var(X-Y) = Var(X) + Var(Y) - 2\,Cov(X,Y)$$

$$16 = 4 + 9 - 2\,Cov(X,Y).$$

Hence

$$Cov(X,Y) = -\frac{3}{2}.$$

Remark. The last Theorem can be extended to three or more random variables. In case of three random variables X, Y,Z, we have

$$\begin{aligned} &Var(X+Y+Z) \\ &= Var(X) + Var(Y) + Var(Z) \\ &\qquad + 2Cov(X,Y) + 2Cov(Y,Z) + 2Cov(Z,X). \end{aligned}$$

To see this consider

$$
\begin{aligned}
&Var(X+Y+Z)\\
&= Var((X+Y)+Z)\\
&= Var(X+Y)+Var(Z)+2Cov(X+Y,Z)\\
&= Var(X+Y)+Var(Z)+2Cov(X,Z)+2Cov(Y,Z)\\
&= Var(X)+Var(Y)+2Cov(X,Y)\\
&\qquad +Var(Z)+2Cov(X,Z)+2Cov(Y,Z)\\
&= Var(X)+Var(Y)+Var(Z)\\
&\qquad +2Cov(X,Y)+2Cov(Y,Z)+2Cov(Z,X).
\end{aligned}
$$

Theorem. If X and Y are independent random variables with E(X) = 0 = E(Y), then

Var(XY) = Var(X) Var(Y).

Proof:

$$
\begin{aligned}
Var(XY) &= E\left((XY)^2\right) - \left(E(X)\,E(Y)\right)^2 \\
&= E\left((XY)^2\right) \\
&= E\left(X^2 Y^2\right) \\
&= E\left(X^2\right)\,E\left(Y^2\right) \qquad \text{(by independence of } X \text{ and } Y\text{)} \\
&= Var(X)\,Var(Y).
\end{aligned}
$$

Example. Let X and Y be independent random variables, each with density

$$
f(x) = \begin{cases} \frac{1}{2\theta} & \text{for } -\theta < x < \theta \\ 0 & \text{otherwise.} \end{cases}
$$

If the $Var(XY) = \frac{64}{9}$, then what is the value of θ ?

Answer:

$$E(X) = \int_{-\theta}^{\theta} \frac{1}{2\theta}\, x\, dx = \frac{1}{2\theta} \left[\frac{x^2}{2}\right]_{-\theta}^{\theta} = 0.$$

Since Y has the same density, we conclude that $E(Y) = 0$. Hence

$$\begin{aligned}
\frac{64}{9} &= Var(XY) \\
&= Var(X)\, Var(Y) \\
&= \left(\int_{-\theta}^{\theta} \frac{1}{2\theta}\, x^2\, dx\right) \left(\int_{-\theta}^{\theta} \frac{1}{2\theta}\, y^2\, dy\right) \\
&= \left(\frac{\theta^2}{3}\right) \left(\frac{\theta^2}{3}\right) \\
&= \frac{\theta^4}{9}.
\end{aligned}$$

Hence, we obtain

$$\theta^4 = 64 \qquad \text{or} \qquad \theta = 2\sqrt{2}.$$

Correlation and Independence

The functional dependency of the random variable Y on the random variable X can be obtained by examining the correlation coefficient. The definition of the correlation coefficient ρ between X and Y is given below.

Definition Let X and Y be two random variables with variances σ_X^2 and σ_Y^2, respectively. Let the covariance of X and Y be $Cov(X,Y)$. Then the correlation coefficient ρ between X and Y is given by

$$\rho = \frac{Cov(X,Y)}{\sigma_X\,\sigma_Y}.$$

Theorem. If X and Y are independent, the correlation coefficient between X and Y is zero.

Proof:

$$\begin{aligned}\rho &= \frac{Cov(X,Y)}{\sigma_X\,\sigma_Y}\\ &= \frac{0}{\sigma_X\,\sigma_Y}\\ &= 0.\end{aligned}$$

Remark. The converse of this theorem is not true. If the correlation coefficient of X and Y is zero, then X and Y are said to be uncorrelated.

Theorem :For any random variables X and Y , the correlation coefficient ρ satisfies

$$-1 \leq \rho \leq 1,$$

Moment Generating Functions

Definition. Let X and Y be two random variables with joint density function f(x, y). A real valued function $M : \mathbb{R}^2 \rightarrow \mathbb{R}$ defined by

$$M(s, t) = E\left(e^{sX+tY}\right)$$

It is easy to see from this definition that

$$M(s,0) = E\left(e^{sX}\right)$$

and

$$M(0,t) = E\left(e^{tY}\right).$$

From this we see that

$$E(X^k) = \left.\frac{\partial^k M(s,t)}{\partial s^k}\right|_{(0,0)}, \qquad E(Y^k) = \left.\frac{\partial^k M(s,t)}{\partial t^k}\right|_{(0,0)},$$

for $k = 1, 2, 3, 4, \ldots$; and

$$E(XY) = \left.\frac{\partial^2 M(s,t)}{\partial s\,\partial t}\right|_{(0,0)}.$$

Example. Let the random variables X and Y have the joint density

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

What is the joint moment generating function for X and Y ?

Answer: The joint moment generating function of X and Y is given by

$$\begin{aligned} M(s,t) &= E\left(e^{sX+tY}\right) \\ &= \int_0^\infty \int_0^\infty e^{sx+ty}\, f(x,y)\, dy\, dx \\ &= \int_0^\infty \int_x^\infty e^{sx+ty}\, e^{-y}\, dy\, dx \\ &= \int_0^\infty \left[\int_x^\infty e^{sx+ty-y}\, dy \right] dx \\ &= \frac{1}{(1-s-t)\,(1-t)}, \qquad \text{provided} \quad s+t<1 \quad \text{and} \quad t<1. \end{aligned}$$

Example :If the joint moment generating function of the random variables X and Y is

$$M(s,t) = e^{(s+3t+2s^2+18t^2+12st)}$$

what is the covariance of X and Y ?

Answer:

$$M(s,t) = e^{(s+3t+2s^2+18t^2+12st)}$$

$$\frac{\partial M}{\partial s} = (1+4s+12t)\,M(s,t)$$

$$\left.\frac{\partial M}{\partial s}\right|_{(0,0)} = 1\,M(0,0)$$

$$= 1.$$

$$\Longrightarrow$$

$$\frac{\partial M}{\partial t} = (3+36t+12s)\,M(s,t)$$

$$\left.\frac{\partial M}{\partial t}\right|_{(0,0)} = 3\,M(0,0)$$

$$= 3.$$

Hence

$$\mu_X = 1 \qquad \text{and} \qquad \mu_Y = 3.$$

Now we compute the product moment of X and Y.

$$\begin{aligned}
\frac{\partial^2 M(s,t)}{\partial s\, \partial t} &= \frac{\partial}{\partial t}\left(\frac{\partial M}{\partial s}\right) \\
&= \frac{\partial}{\partial t}\left(M(s,t)\,(1+4s+12t)\right) \\
&= (1+4s+12t)\,\frac{\partial M}{\partial t} + M(s,t)\,(12).
\end{aligned}$$

Therefore

$$\left.\frac{\partial^2 M(s,t)}{\partial s\, \partial t}\right|_{(0,0)} = 1\,(3) + 1\,(12).$$

Thus

$$E(XY) = 15$$

and the covariance of X and Y is given by

$$\begin{aligned} Cov(X, Y) &= E(XY) - E(X)\, E(Y) \\ &= 15 - (3)\,(1) \\ &= 12. \end{aligned}$$

Theorem. If X and Y are independent and let a and b real parameters then

$$M_{aX+bY}(t) = M_X(at)\, M_Y(bt),$$

Proof: Let $W = aX + bY$. Hence

$$\begin{aligned} M_{aX+bY}(t) &= M_W(t) \\ &= E\left(e^{tW}\right) \\ &= E\left(e^{t(aX+bY)}\right) \\ &= E\left(e^{taX}\, e^{tbY}\right) \\ &= E\left(e^{taX}\right) E\left(e^{tbY}\right) \\ &= M_X(at)\, M_Y(bt). \end{aligned}$$

This theorem is very powerful. It helps us to find the distribution of a linear combination of independent random variables. The following examples illustrate how one can use this theorem to determine distribution of a linear combination.

Exercises

1. Suppose that X_1 and X_2 are random variables with zero mean and unit variance. If the correlation coefficient of X_1 and X_2 is -0.5, then what is the variance of $Y = \sum_{k=1}^{2} k^2 X_k$?

2. If the joint density of the random variables X and Y is

$$f(x,y) = \begin{cases} \frac{1}{8} & \text{if } (x,y) \in \{\, (x,0),\ (0,-y) \,|\, x,y = -2,-1,1,2 \,\} \\ 0 & \text{otherwise,} \end{cases}$$

what is the covariance of X and Y ? Are X and Y independent?

3. Suppose the random variables X and Y are independent and identically distributed. Let $Z = aX + Y$. If the correlation coefficient between X and Z is $\frac{1}{3}$, then what is the value of the constant a ?

7. If the joint probability density function of X and Y is

$$f(x,y) = \begin{cases} 1 & \text{if } 0 < x < 1;\ 0 < y < 1 \\ \\ 0 & \text{elsewhere,} \end{cases}$$

then what is the joint moment generating function of X and Y ?

8. Let the joint density function of X and Y be

$$f(x,y) = \begin{cases} \frac{1}{36} & \text{if } 1 \le x = y \le 6 \\ \\ \frac{2}{36} & \text{if } 1 \le x < y \le 6. \end{cases}$$

What is the correlation coefficient of X and Y ?

14. Let Y and Z be two random variables. If $Var(Y) = 4$, $Var(Z) = 16$, and $Cov(Y, Z) = 2$, then what is $Var(3Z - 2Y)$?

7 The joint density function of two continuous random variables X and Y is

$$f(x, y) = \begin{cases} cxy & 0 < x < 4, 1 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the value of the constant c.

(b) Find $P(1 < X < 2, 2 < Y < 3)$.

(c) Find $P(X \geq 3, Y \leq 2)$.

8

The joint probability function of two discrete random variables X and Y is given by $f(x, y) = c(2x + y)$, where x and y can assume all integers such that $0 \leq x \leq 2$, $0 \leq y \leq 3$, and $f(x, y) = 0$ otherwise.

(a) Find the value of the constant c.

(b) Find $P(X = 2, Y = 1)$.

(c) Find $P(X \geq 1, Y \leq 2)$.

See you next Lecture

$$M(s,t) = E\left(e^{sx+ty}\right)$$

$$= \int_0^1 \int_0^1 e^{sx+ty}\, dy\, dx$$

$$= \int_0^1 t e^{sx+ty} \Big|_0^1 dx$$

$$= \frac{1}{t} \int_0^1 e^{sx+t}\, dx$$

$$= \frac{1}{t} \left[s e^{sx+t} \Big|_0^1 \right.$$

$$= st e^{s}$$

LECTURE 19#

Miscellaneous Problems

1) Find the expectation of the sum of points in tossing a pair of fair dice.

2) Suppose that the random variables X and Y have a joint density function given by

$$f(x, y) = \begin{cases} c(2x + y) & 2 < x < 6, 0 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

Find (a) the constant c, (b) the marginal distribution functions for X and Y, (c) the marginal density functions for X and Y, (d) $P(3 < X < 4, Y > 2)$, (e) $P(X > 3)$, (f) $P(X + Y > 4)$, (g) the joint distribution function, (h) whether X and Y are independent.

3) If X and Y have the joint density function

$$f(x, y) = \begin{cases} \frac{3}{4} + xy & 0 < x < 1, 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

find (a) $f(y \mid x)$, (b) $P(Y > \frac{1}{2} \mid \frac{1}{2} < X < \frac{1}{2} + dx)$.

4)

The joint density function of the random variables X and Y is given by

$$f(x, y) = \begin{cases} 8xy & 0 \leq x \leq 1, 0 \leq y \leq x \\ 0 & \text{otherwise} \end{cases}$$

Find (a) the marginal density of X, (b) the marginal density of Y, (c) the conditional density of X, (d) the conditional density of Y.

5)

The distribution function for a random variable X is

$$F(x) = \begin{cases} 1 - e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Find (a) the density function, (b) the probability that $X > 2$, and (c) the probability that $-3 < X \leq 4$.

6)

A random variable X has the density function $f(x) = c/(x^2 + 1)$, where $-\infty < x < \infty$. (*a*) Find the value of the constant c. (b) Find the probability that X^2 lies between $1/3$ and 1.

7) Find (a) the variance, (b) the standard deviation of the sum obtained in tossing a pair of fair dice .

8) Find the characteristic function of the random variable X having density function given by

$$f(x) = \begin{cases} 1/2a & |x| < a \\ 0 & \text{otherwise} \end{cases}$$

9) A random variable X has density function given by

$$f(x) = \begin{cases} 2e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Find (a) the moment generating function, (b) the first four moments about the origin.

10) For the random variable of Problem 9.

(a) find $P(|X - \mu| > 1)$.

b) Use Chebyshev's inequality to obtain an upper bound on $P(|X - \mu| > 1)$ and compare with the result in (a).

LECTURE 20#

Miscellaneous Problems

1) Find the expectation of the sum of points in tossing a pair of fair dice.

Let X and Y be the points showing on the two dice. We have

$$E(X) = E(Y) = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + \cdots + 6\left(\frac{1}{6}\right) = \frac{7}{2}$$

Then, by Theorem 3-2,

$$E(X + Y) = E(X) + E(Y) = 7$$

2.33. Suppose that the random variables X and Y have a joint density function given by

$$f(x, y) = \begin{cases} c(2x + y) & 2 < x < 6, 0 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

Find (a) the constant c, (b) the marginal distribution functions for X and Y, (c) the marginal density functions for X and Y, (d) $P(3 < X < 4, Y > 2)$, (e) $P(X > 3)$, (f) $P(X + Y > 4)$, (g) the joint distribution function, (h) whether X and Y are independent.

(a) The total probability is given by

$$\int_{x=2}^{6}\int_{y=0}^{5} c(2x + y)\,dx\,dy = \int_{x=2}^{6} c\left(2xy + \frac{y^2}{2}\right)\Bigg|_0^5 dx$$

$$= \int_{x=2}^{6} c\left(10x + \frac{25}{2}\right)dx = 210c$$

For this to equal 1, we must have $c = 1/210$.

(b) The marginal distribution function for X is

$$F_1(x) = P(X \le x) = \int_{u=-\infty}^{x}\int_{v=-\infty}^{\infty} f(u, v)\,du\,dv$$

$$= \begin{cases} \displaystyle\int_{u=-\infty}^{x}\int_{v=-\infty}^{\infty} 0\,du\,dv = 0 & x < 2 \\ \displaystyle\int_{u=2}^{x}\int_{v=0}^{5} \frac{2u + v}{210}\,du\,dv = \frac{2x^2 + 5x - 18}{84} & 2 \le x < 6 \\ \displaystyle\int_{u=2}^{6}\int_{v=0}^{5} \frac{2u + v}{210}\,du\,dv = 1 & x \ge 6 \end{cases}$$

If X and Y have the joint density function

$$f(x, y) = \begin{cases} \frac{3}{4} + xy & 0 < x < 1, 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

find (a) $f(y \mid x)$, (b) $P(Y > \frac{1}{2} \mid \frac{1}{2} < X < \frac{1}{2} + dx)$.

(a) For $0 < x < 1$,

$$f_1(x) = \int_0^1 \left(\frac{3}{4} + xy\right) dy = \frac{3}{4} + \frac{x}{2}$$

and
$$f(y \mid x) = \frac{f(x, y)}{f_1(x)} = \begin{cases} \dfrac{3 + 4xy}{3 + 2x} & 0 < y < 1 \\ 0 & \text{other } y \end{cases}$$

For other values of x, $f(y \mid x)$ is not defined.

(b)
$$P(Y > \tfrac{1}{2} \mid \tfrac{1}{2} < X < \tfrac{1}{2} + dx) = \int_{1/2}^{\infty} f(y \mid \tfrac{1}{2})\, dy = \int_{1/2}^{1} \frac{3 + 2y}{4}\, dy = \frac{9}{16}$$

The marginal distribution function for Y is

$$F_2(y) = P(Y \leq y) = \int_{u=-\infty}^{\infty}\int_{v=-\infty}^{y} f(u, v)\,du\,dv$$

$$= \begin{cases} \displaystyle\int_{u=-\infty}^{\infty}\int_{v=-8}^{y} 0\,du\,dv = 0 & y < 0 \\ \displaystyle\int_{u=0}^{6}\int_{v=0}^{y} \frac{2u + v}{210}\,du\,dv = \frac{y^2 + 16y}{105} & 0 \leq y < 5 \\ \displaystyle\int_{u=2}^{6}\int_{v=0}^{5} \frac{2u + v}{210}\,du\,dv = 1 & y \geq 5 \end{cases}$$

(c) The marginal density function for X is, from part (b),

$$f_1(x) = \frac{d}{dx}F_1(x) = \begin{cases} (4x + 5)/84 & 2 < x < 6 \\ 0 & \text{otherwise} \end{cases}$$

The marginal density function for Y is, from part (b),

$$f_2(y) = \frac{d}{dy}F_2(y) = \begin{cases} (2y + 16)/105 & 0 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

(d) $$P(3 < X < 4, Y > 2) = \frac{1}{210}\int_{x=3}^{4}\int_{y=2}^{5}(2x + y)\,dx\,dy = \frac{3}{20}$$

(e) $$P(X > 3) = \frac{1}{210}\int_{x=3}^{6}\int_{y=0}^{5}(2x + y)\,dx\,dy = \frac{23}{28}$$

(f)
$$P(X + Y > 4) = \iint\limits_{\mathscr{R}} f(x, y)\,dx\,dy$$

where $\mathscr{R}$ is the shaded region of Fig. 2-20. Although this can be found, it is easier to use the fact that

$$P(X + Y > 4) = 1 - P(X + Y \leq 4) = 1 - \iint\limits_{\mathscr{R}} f(x, y)\,dx\,dy$$

where $\mathscr{R}'$ is the cross-hatched region of Fig. 2-20. We have

$$P(X + Y \leq 4) = \frac{1}{210}\int_{x=2}^{4}\int_{y=0}^{4-x}(2x + y)\,dx\,dy = \frac{2}{35}$$

Thus $P(X + Y > 4) = 33/35$.

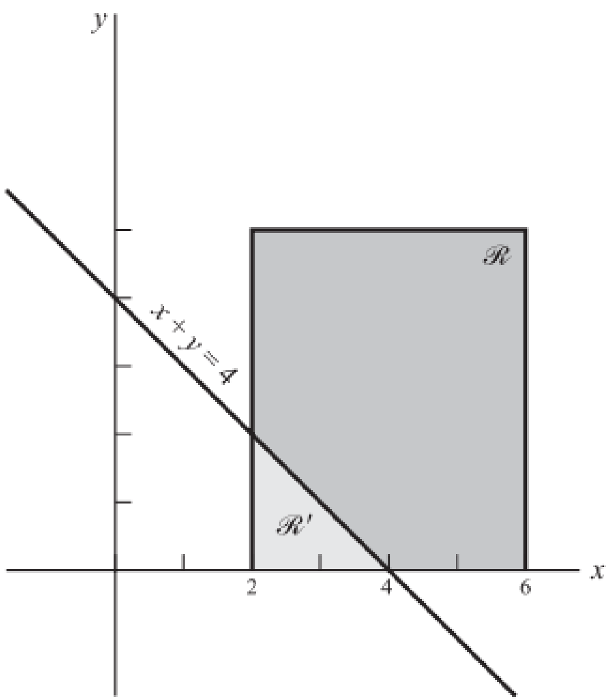


Fig. 2-20

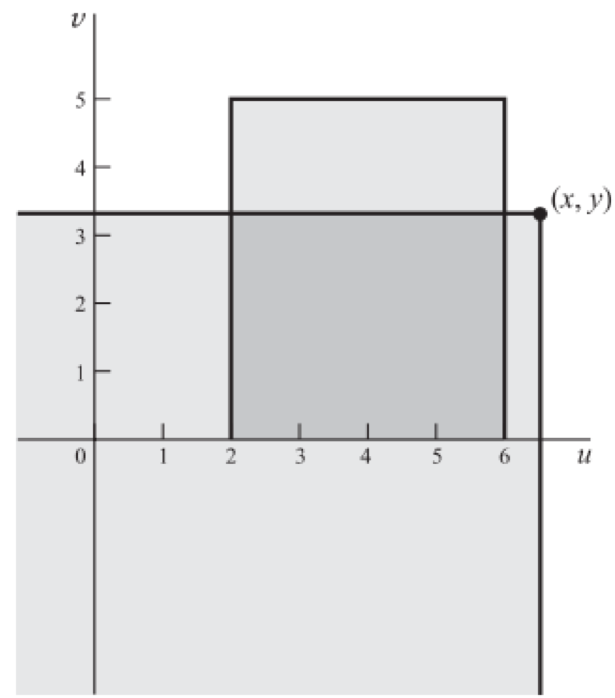


Fig. 2-21

(g) The joint distribution function is

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

In the uv plane (Fig. 2-21) the region of integration is the intersection of the quarter plane $u \leq x$, $v \leq y$ and the rectangle $2 < u < 6$, $0 < v < 5$ [over which $f(u, v)$ is nonzero]. For (x, y) located as in the figure, we have

$$F(x, y) = \int_{u=2}^{6} \int_{v=0}^{y} \frac{2u + v}{210}\, du\, dv = \frac{16y + y^2}{105}$$

When (x, y) lies inside the rectangle, we obtain another expression, etc. The complete results are shown in Fig. 2-22.

(h) The random variables are dependent since

$$f(x, y) \neq f_1(x) f_2(y)$$

or equivalently, $F(x, y) \neq F_1(x) F_2(y)$.

3) **2.28.** If X and Y have the joint density function

$$f(x, y) = \begin{cases} \frac{3}{4} + xy & 0 < x < 1, 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

find (a) $f(y \mid x)$, (b) $P(Y > \frac{1}{2} \mid \frac{1}{2} < X < \frac{1}{2} + dx)$.

(a) For $0 < x < 1$,

$$f_1(x) = \int_0^1 \left(\frac{3}{4} + xy\right) dy = \frac{3}{4} + \frac{x}{2}$$

and

$$f(y \mid x) = \frac{f(x, y)}{f_1(x)} = \begin{cases} \dfrac{3 + 4xy}{3 + 2x} & 0 < y < 1 \\ 0 & \text{other } y \end{cases}$$

For other values of x, $f(y \mid x)$ is not defined.

(b)

$$P(Y > \tfrac{1}{2} \mid \tfrac{1}{2} < X < \tfrac{1}{2} + dx) = \int_{1/2}^{\infty} f(y \mid \tfrac{1}{2})\, dy = \int_{1/2}^{1} \frac{3 + 2y}{4}\, dy = \frac{9}{16}$$

4)

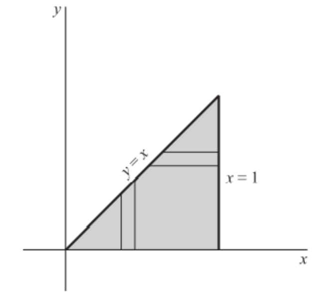


Fig. 2-17

2.29. The joint density function of the random variables X and Y is given by

$$f(x, y) = \begin{cases} 8xy & 0 \le x \le 1, 0 \le y \le x \\ 0 & \text{otherwise} \end{cases}$$

Find (a) the marginal density of X, (b) the marginal density of Y, (c) the conditional density of X, (d) the conditional density of Y.

The region over which $f(x, y)$ is different from zero is shown shaded in Fig. 2-17.

(a) To obtain the marginal density of X, we fix x and integrate with respect to y from 0 to x as indicated by the vertical strip in Fig. 2-17. The result is

$$f_1(x) = \int_{y=0}^{x} 8xy\,dy = 4x^3$$

for $0 < x < 1$. For all other values of x, $f_1(x) = 0$.

(b) Similarly, the marginal density of Y is obtained by fixing y and integrating with respect to x from $x = y$ to $x = 1$, as indicated by the horizontal strip in Fig. 2-17. The result is, for $0 < y < 1$,

$$f_2(y) = \int_{x=y}^{1} 8xy\,dx = 4y(1 - y^2)$$

For all other values of y, $f_2(y) = 0$.

(c) The conditional density function of X is, for $0 < y < 1$,

$$f_1(x \mid y) = \frac{f(x, y)}{f_2(y)} = \begin{cases} 2x/(1 - y^2) & y \le x \le 1 \\ 0 & \text{other } x \end{cases}$$

The conditional density function is not defined when $f_2(y) = 0$.

(d) The conditional density function of Y is, for $0 < x < 1$,

$$f_2(y \mid x) = \frac{f(x, y)}{f_1(x)} = \begin{cases} 2y/x^2 & 0 \le y \le x \\ 0 & \text{other } y \end{cases}$$

The conditional density function is not defined when $f_1(x) = 0$.

Check:

$$\int_0^1 f_1(x)\,dx = \int_0^1 4x^3\,dx = 1, \quad \int_0^1 f_2(y)\,dy = \int_0^1 4y(1 - y^2)\,dy = 1$$

$$\int_y^1 f_1(x\,|\,y)\,dx = \int_y^1 \frac{2x}{1 - y^2}\,dx = 1$$

$$\int_0^x f_2(y\,|\,x)\,dy = \int_0^x \frac{2y}{x^2}\,dy = 1$$

2.7. The distribution function for a random variable X is

$$F(x) = \begin{cases} 1 - e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Find (a) the density function, (b) the probability that $X > 2$, and (c) the probability that $-3 < X \leq 4$.

(a)
$$f(x) = \frac{d}{dx}F(x) = \begin{cases} 2e^{-2x} & x > 0 \\ 0 & x < 0 \end{cases}$$

(b)
$$P(X > 2) = \int_2^{\infty} 2e^{-2u}\,du = -e^{-2u}\Big|_2^{\infty} = e^{-4}$$

Another method

By definition, $P(X \leq 2) = F(2) = 1 - e^{-4}$. Hence,

$$P(X > 2) = 1 - (1 - e^{-4}) = e^{-4}$$

(c)
$$P(-3 < X \leq 4) = \int_{-3}^{4} f(u)\,du = \int_{-3}^{0} 0\,du + \int_0^4 2e^{-2u}\,du$$
$$= -e^{-2u}\Big|_0^4 = 1 - e^{-8}$$

Another method

$$P(-3 < X \leq 4) = P(X \leq 4) - P(X \leq -3)$$
$$= F(4) - F(-3)$$
$$= (1 - e^{-8}) - (0) = 1 - e^{-8}$$

2.5. A random variable X has the density function $f(x) = c/(x^2 + 1)$, where $-\infty < x < \infty$. (*a*) Find the value of the constant c. (b) Find the probability that X^2 lies between $1/3$ and 1.

(a) We must have $\int_{-\infty}^{\infty} f(x)\,dx = 1$, i.e.,

$$\int_{-\infty}^{\infty} \frac{c\,dx}{x^2 + 1} = c \tan^{-1} x \Big|_{-\infty}^{\infty} = c\left[\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right] = 1$$

so that $c = 1/\pi$.

(b) If $\frac{1}{3} \le X^2 \le 1$, then either $\frac{\sqrt{3}}{3} \le X \le 1$ or $-1 \le X \le -\frac{\sqrt{3}}{3}$. Thus the required probability is

$$\frac{1}{\pi}\int_{-1}^{-\sqrt{3}/3} \frac{dx}{x^2 + 1} + \frac{1}{\pi}\int_{\sqrt{3}/3}^{1} \frac{dx}{x^2 + 1} = \frac{2}{\pi}\int_{\sqrt{3}/3}^{1} \frac{dx}{x^2 + 1}$$

$$= \frac{2}{\pi}\left[\tan^{-1}(1) - \tan^{-1}\left(\frac{\sqrt{3}}{3}\right)\right]$$

$$= \frac{2}{\pi}\left(\frac{\pi}{4} - \frac{\pi}{6}\right) = \frac{1}{6}$$

3.8. Find (a) the variance, (b) the standard deviation of the sum obtained in tossing a pair of fair dice.

(a) Referring to Problem 3.2, we have $E(X) = E(Y) = 1/2$. Moreover,

$$E(X^2) = E(Y^2) = 1^2\left(\frac{1}{6}\right) + 2^2\left(\frac{1}{6}\right) + \cdots + 6^2\left(\frac{1}{6}\right) = \frac{91}{6}$$

Then, by Theorem 3-4,

$$\mathrm{Var}(X) = \mathrm{Var}(Y) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

and, since X and Y are independent, Theorem 3-7 gives

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) = \frac{35}{6}$$

(b) $$\sigma_{X+Y} = \sqrt{\mathrm{Var}(X + Y)} = \sqrt{\frac{35}{6}}$$

3.21. Find the characteristic function of the random variable X having density function given by

$$f(x) = \begin{cases} 1/2a & |x| < a \\ 0 & \text{otherwise} \end{cases}$$

The characteristic function is given by

$$E(e^{i\omega X}) = \int_{-\infty}^{\infty} e^{i\omega x} f(x)\,dx = \frac{1}{2a}\int_{-a}^{a} e^{i\omega x}\,dx$$

$$= \frac{1}{2a}\,\frac{e^{i\omega x}}{i\omega}\bigg|_{-a}^{a} = \frac{e^{ia\omega} - e^{-ia\omega}}{2ia\omega} = \frac{\sin a\omega}{a\omega}$$

using Euler's formulas (see Problem 3.20) with $\theta = a\omega$.

3.18. A random variable X has density function given by

$$f(x) = \begin{cases} 2e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Find (a) the moment generating function, (b) the first four moments about the origin.

(a)
$$M(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$$

$$= \int_{0}^{\infty} e^{tx}(2e^{-2x})\,dx = 2\int_{0}^{\infty} e^{(t-2)x}\,dx$$

$$= \left.\frac{2e^{(t-2)x}}{t-2}\right|_{0}^{\infty} = \frac{2}{2-t}, \quad \text{assuming } t < 2$$

(b) If $|t| < 2$ we have

$$\frac{2}{2-t} = \frac{1}{1-t/2} = 1 + \frac{t}{2} + \frac{t^2}{4} + \frac{t^3}{8} + \frac{t^4}{16} + \cdots$$

But
$$M(t) = 1 + \mu t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \mu_4' \frac{t^4}{4!} + \cdots$$

Therefore, on comparing terms, $\mu = \frac{1}{2}$, $\mu_2' = \frac{1}{2}$, $\mu_3' = \frac{3}{4}$, $\mu_4' = \frac{3}{2}$.

3.31. For the random variable of Problem 3.18, (a) find $P(|X - \mu| > 1)$. (b) Use Chebyshev's inequality to obtain an upper bound on $P(|X - \mu| > 1)$ and compare with the result in (a).

(a) From Problem 3.18, $\mu = 1/2$. Then

$$P(|X - \mu| < 1) = P\left(\left|X - \frac{1}{2}\right| < 1\right) = P\left(-\frac{1}{2} < X < \frac{3}{2}\right)$$

$$= \int_0^{3/2} 2e^{-2x}\,dx = 1 - e^{-3}$$

Therefore $$P\left(\left|X - \frac{1}{2}\right| \geq 1\right) = 1 - (1 - e^{-3}) = e^{-3} = 0.04979$$

(b) From Problem 3.18, $\sigma^2 = \mu_2' - \mu^2 = 1/4$. Chebyshev's inequality with $\epsilon = 1$ then gives

$$P(|X - \mu| \geq 1) \leq \sigma^2 = 0.25$$

Comparing with (a), we see that the bound furnished by Chebyshev's inequality is here quite crude. In practice, Chebyshev's inequality is used to provide estimates when it is inconvenient or impossible to obtain exact values.