Lecture #2

Review of Probability Theory and Random Variables

1-The Axioms of Probability

Axiom 1. For every event A in given class, $P(A) \geq 0$

Axiom 2. For the sure or certain event S in the class, $P(S) = 1$

Axiom 3. For any number of mutually exclusive events A1, A2, …, in the class,

$$P(A1 \cup A2 \cup \dots) = P(A1) + P(A2) + \dots$$

In particular, for two mutually exclusive events A1 and A2, $P(A1 \cup A2) = P(A1) + P(A2)$

2- Some Important Probability Theorems

The following theorems on probability are important.

1) For every event A, $0 \leq P(A) \leq 1$, i.e., a probability between 0 and 1.
2) For $\emptyset$, the empty set, $P(\emptyset) = 0$, the impossible event has probability zero.
3) If A' is the complement of A, then $P(A') = 1 - P(A)$
4) If A and B are any two events, then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

3- Independent Events

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

A simple example: if one toss two fair coins, the event T or H on first coin is independent on that occurred on the second coin.

4- Conditional Probability

Let A and B be two events such that $P(A) > 0$. Denote $P(B \mid A)$ the probability of B given that A has occurred. $P(B|A) = \frac{P(A \cap B)}{P(A)}$

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

In words, this is saying that the probability that both A and B occur is equal to the probability that A occurs times the probability that B occurs given that A has occurred. We call P(B | A) the conditional probability of B given A, i.e., the probability that B will occur given that A has occurred. It is easy to show that conditional probability satisfies the axioms of probability previously discussed.

Similarly,

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

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

- If A and B are mutually exclusive events then $P(A \cap B) = 0$, hence both $P(A|B)$ and $P(B|A)$ are zero.
- If A and B are independent event, $P(A \cap B) = P(A).P(B)$, then $P(A|B) = P(A)$ and $P(B|A) = P(B)$.

5- Outcomes as Elements of Events

For some applications the event may contains few outcomes. This like considering the event as a set that containing elements.

Let X = { x_1 , x_2 , x_3 , ...x_N} and Y= {y_1 , y_2 , y_3 , ...y_M}

Then all the above relations and theorems will be applied to the outcomes x_i and y_j .

Example: Binary Symmetric Channel

In binary symmetric channel, the transmitter (T) sends binary values (0_T or 1_T), and the receiver (R) receives binary values (0_R or 1_R). If the channel is perfect or noiseless then the receiver will receive exactly what the transmitter sends and in this case:

We have two objects the transmitter (T) and the receiver (R) , with

$$T = \{ 0_T \ , \ 1_T \} \quad \text{and} \quad R = \{ 0_R \ , \ 1_R \}$$

Noiseless Channel: $P(0_R) = P(0_T)$ and $P(1_R) = P(1_T)$ with

$P(0_R \mid 1_T) = P(1_R \mid 0_T) = 0$ (no transition error)

Also $P(0_R \mid 0_T) = P(1_R \mid 1_T) = 1$ (totally correct transition)

In practice the channel is not noiseless (i.e. noisy).

Noisy Channel: $P(0_R) \neq P(0_T)$ and $P(1_R) \neq P(1_T)$ with

$P(0_R \mid 1_T) = P(1_R \mid 0_T) \neq 0$ (there is transition error)

Also $P(0_R \mid 0_T) = P(1_R \mid 1_T) \neq 1$ (partially correct transition)

Let us consider the noisy case:

Given: $P(0_T) = 0.6$, $P(1_T) = 0.4$ and $P(0_R \mid 1_T) = P(1_R \mid 0_T) = 0.1$

Find: $P(0_R \mid 0_T)$, $P(1_R \mid 1_T)$, all $p(T \cap R)$, $P(0_R)$ and $P(1_R)$

Solution:

Since $P(0_R \mid 1_T) + P(0_R \mid 0_T) = 1$ then $P(0_R \mid 0_T) = 1 - 0.1 = 0.9$

Similarly $P(1_R \mid 0_T) + P(1_R \mid 1_T) = 1$ then $P(1_R \mid 1_T) = 1 - 0.1 = 0.9$

Using the relation $p(T \cap R) = P(R|T).\ P(T)$ for all T and R values gives

$P(0_R \cap 0_T) = P(0_R \mid 0_T)\ P(0_T) = 0.9 \text{x} 0.6 = 0.54$

$P(1_R \cap 1_T) = P(1_R \mid 1_T)\ P(1_T) = 0.9 \text{x} 0.4 = 0.36$

$P(0_R \cap 1_T) = P(0_R \mid 1_T)\ P(1_T) = 0.1 \text{x} 0.4 = 0.04$

$P(1_R \cap 0_T) = P(1_R \mid 0_T)\ P(0_T) = 0.1 \text{x} 0.6 = 0.06$

Then find the receiver probabilities by;

$P(0_R) = P(0_R \cap 0_T) + P(0_R \cap 1_T) = 0.54 + 0.04 = 0.58$

$P(1_R) = P(1_R \cap 0_T) + P(1_R \cap 1_T) = 0.36 + 0.06 = 0.42$

It is clear to see that sum of P(T) is one and so also the sum of P(R).

6- Types of Random Variables

A. Discrete Random Variables (DRV)

Here we have limited number of values for the variables. We call this number as N. Each value of DRV has probability value p_i usually determined by probability rules. The probability of DRV may be defined by the relative frequency $p_i = \frac{n_i}{N}$.

X = {x_1 , x_2 , x_3 , x_4 , x_5 , …… x_N } , **with** $\sum_{i=1}^{N} p_i$ **=1**

P = {p_1 , p_2 , p_3 , p_4 , p_5 , …… p_N } , $p_i \neq 0$ *in general*

Example: In die experiment the output number is DRV. The number of variables is limited to 6, thus it is discrete random variable.

X	x_1	x_2	x_3	x_4	x_5	x_6
x_i (RV)	1	2	3	4	5	6
p_i	1/6	1/6	1/6	1/6	1/6	1/6

The above can be considered as a set of 6 elements, each with probability of 1/6.

B. Continues Random Variables (CRV)

In CRV, the random variable has unlimited values (N → ∞). Thus p_i is undefined this case.

Example: Noise waveform

The above waveform does not have fixed number of voltage values (i.e unlimited or ∞). In this case we can not define certain probability of such voltage. Instead, we may talk about range of voltage, maximum, minimum, average….etc.

Probability Density Function (pdf) of CRV

Probability Density Function (or pdf f(x)) is defined and used for CRV to determine probability of given range of the CRV x as given below:

$$p(a < x < b) = \int_a^b f(x)dx \qquad \text{and} \qquad p(x=a) =0 \ \ (\text{why})$$

Since f(x) is density function then its unit is probability divided by unit of x. After integration the unit will be the probability.

Properties of pdf f(x)

a) f(x) is non-negative function (may be zero or +ve)
b) f(x) < 1
c) $\int_{-\infty}^{\infty} f(x)dx = 1$

Any function satisfies the above properties is pdf and vice versa.

7-Expectations and Moments

Expectations

Expectations or averages are the mean value of the function with the random variable being the independent variable of the function say h(x). The general expression for expectation is given by:

For DRV $\quad E[h(x)] =\overline{h(x)}= \sum_{x_i} h(x_i).P(x_i)$

For CRV $\quad E[h(x)]= \overline{h(x)}= \int_{-\infty}^{\infty} h(x).f(x)dx$

Properties of Expectation:

a) E[h(x)] may be –ve, 0, or +ve.
b) For c=constant E[c.h(x)] = c.E[h(x)] , and E[c] =c
c) E[h(x)] is linear operation: $E[a.h_1(x) + b.h_2(x)] = a.E[h_1(x)] + b.E[h_2(x)]$ (a & b are constants)
d) E[h(x)] is non-negative when h(x) is positive function as in $h(x)=x^2$ for example.

Moments

The moments are expectations of special functions of the random variables. There are two types of moments:

a) Moments about the origin $m_k=E[x^k]$
b) Moments about the mean is $M_k=E[(x-\overline{x})^k]$

Some useful moments:

$m_1 = E[x] = \overline{x}$ (this is the mean value or DC level of the RV x)

$m_2 = E[x^2] = \overline{x^2}$ (the mean square value or the average total power of the RV x)

$M_1 = 0$ (why)

$M_2 = E[(x-\overline{x})^2] = \sigma^2$ = variance , σ = standard deviation of RV x.

Important relation: $\sigma^2 = \overline{x^2} - \overline{x}^2$

Proof:

$$\text{L.H.S} = \sigma^2 = M_2 = E[(x-\overline{x})^2] = E[x^2 - 2x\overline{x} + \overline{x}^2] = E[x^2] - E[2x\overline{x}] + E[\overline{x}^2]$$

$$= \overline{x^2} - 2.\overline{x}\,E[x] + \overline{x}^2 = \overline{x^2} - 2.\overline{x}\,\overline{x} + \overline{x}^2 = \overline{x^2} - 2.\overline{x}^2 + \overline{x}^2$$

$$= \overline{x^2} - \overline{x}^2 = \text{R.H.S}$$

8- Joint Distribution

When we have two random variables in the same experiment or expression the resultant 2-variable distribution is called Joint Distribution or Joint RV. There also exist joint distributions involving more than two random variables.

Joint DRVs

Consider the following two sets of DRV;

$X = \{x_1, x_2, x_3, x_4, x_5, \ldots\ldots x_N\}$

$Y = \{y_1, y_2, y_3, y_4, y_5, \ldots\ldots y_M\}$

The joint Probability of x_i and y_j = $P(x_i \cap y_j) = P(x_i \text{ and } y_j) = P(x_i , y_j)$

with $\sum_{i=1}^{N} P(x_i) = 1$, $\sum_{j=1}^{M} P(y_j) = 1$, $\sum_x \sum_y P(x_i, y_j) = 1$

Also $P(x_i , y_j) = P(x_i).P(y_j)$ (for independent RV)

$P(x_i , y_j) = 0$ (for mutual exclusive RV)

$$P(x_i) = \sum_y P(x_i, y_j) \quad , \quad P(y_j) = \sum_x P(x_i, y_j)$$

Joint CRVs

Here we have joint continues random variable.

$f(x)$ = pdf of x then $\int_{-\infty}^{\infty} f(x)dx = 1$

$f(y)$ = pdf of y then $\int_{-\infty}^{\infty} f(y)dy = 1$

$f(x,y)$ = joint pdf of x and y then $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$

$f(x,y)= f(x).f(y)$ (for independent RV)

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

Example-1 Discrete Random Variable

Let the random variable x_i is defined as the absolute value of the difference of the two numbers occurring in the experiment of the throwing two dice. Then:

a) Find all possible values of x_i with their probabilities (this also known as the sample space of the experiment)

b) Find : $\overline{x}$, $\overline{x^2}$, σ^2 , $E[4x^2]$, $E[(x-2)^2]$, $E[e^x]$

Solution:

a) Following is the table of all possible outcomes of the experiment with **x=|A-B|**

A \ B	1	2	3	4	5	6
1	**0**	**1**	**2**	**3**	**4**	**5**
2	**1**	**0**	**1**	**2**	**3**	**4**
3	**2**	**1**	**0**	**1**	**2**	**3**
4	**3**	**2**	**1**	**0**	**1**	**2**
5	**4**	**3**	**2**	**1**	**0**	**1**
6	**5**	**4**	**3**	**2**	**1**	**0**

Thus,

X	x_1	x_2	x_3	x_4	x_5	x_6
x_i	0	1	2	3	4	5
p_i	6/36	10/36	8/36	6/36	4/36	2/36

It is important to check that $\sum_{i=1}^{N} P(x_i) = 1$

b) Find : $\overline{x}$, $\overline{x^2}$, σ^2 , $E[4x^2]$, $E[(x-2)^2]$, $E[e^{-x}]$

$\overline{x} = \sum_{x_i} x_i.P(x_i) = 0x\frac{6}{36} + 1x\frac{10}{36} + 2x\frac{8}{36} + 3x\frac{6}{36} + 4x\frac{4}{36} + 5x\frac{2}{36}$ =70/36=1.9444

$\overline{x^2} = \sum_{x_i} x_i^2 . P(x_i) = 0^2 \text{x} \frac{6}{36} + 1^2 \text{x} \frac{10}{36} + 2^2 \text{x} \frac{8}{36} + 3^2 \text{x} \frac{6}{36} + 4^2 \text{x} \frac{4}{36} + 5^2 \text{x} \frac{2}{36}$

$= 0\text{x}\frac{6}{36} + 1\text{x}\frac{10}{36} + 4\text{x}\frac{8}{36} + 9\text{x}\frac{6}{36} + 16\text{x}\frac{4}{36} + 25\text{x}\frac{2}{36}$ = 210/36 = 5.833

$\sigma^2 = \overline{x^2} - \overline{x}^2$ = 5.833- $(1.944)^2$ = 2.054

$E[4x^2] = 4.E[x^2] = 4.\ \overline{x^2}$ = 4 x 5.833 = 23.332

$E[(x-2)^2] = E[x^2 -2x +4] = \overline{x^2} -2.\ \overline{x} +4$ = 5.833 -2x1.9444 + 4 = 5.944

$E[e^x] = \sum_{x_i} e^{x_i} . P(x_i) = e^0 \text{x} \frac{6}{36} + e^1 \text{x} \frac{10}{36} + e^2 \text{x} \frac{8}{36} + e^3 \text{x} \frac{6}{36} + e^4 \text{x} \frac{4}{36} + e^5 \text{x} \frac{2}{36}$

$= 1\text{x}\frac{6}{36} + 2.718\text{x}\frac{10}{36} + 7.39\text{x}\frac{8}{36} + 20.085\text{x}\frac{6}{36} + 54.6\text{x}\frac{4}{36} + 148.4\text{x}\frac{2}{36} = 20.22$

Q-For Example-1 find P(x>2) , P(1 < x < 4), E[1-x]

Example-2 Continues Random Variable

Consider CRV having the following pdf ;

f(x) = Ax for 0 < x < 2 and f(x)=0 elsewhere

a) Find the constant A
b) Find P(x > 1) , P(0.5 < x < 1) , $\overline{x}$, $\overline{x^2}$, σ^2

Solution:

a) Using the fact that $\int_{-\infty}^{\infty} f(x)dx = 1$

Then $\int_0^2 Ax.dx = 1$ gives 2A=1 thus, A=1/2 and the pdf can be written as:

$$f(x) = \begin{cases} \frac{x}{2} & 0 < x < 2 \\ 0 & elsewhere \end{cases}$$

b) P(x > 1) = $\int_1^2 f(x)dx = \int_1^2 \frac{x}{2} dx = \left[\frac{x^2}{4}\right]_1^2$ = ¾

P(0.5 < x < 1) = $\int_{0.5}^1 f(x)dx = \int_{0.5}^1 \frac{x}{2} dx = \left[\frac{x^2}{4}\right]_{0.5}^1 = 0.1875$

$$\overline{x} = \int_{-\infty}^{\infty} x.f(x)dx = \frac{1}{2}\int_0^2 x^2 dx = \left[\frac{x^3}{6}\right]_0^2 = \frac{8}{6} = \frac{4}{3}$$

$$\overline{x^2} = \int_{-\infty}^{\infty} x^2.f(x)dx = \frac{1}{2}\int_0^2 x^3 dx = \left[\frac{x^4}{8}\right]_0^2 = \frac{16}{8} = 2$$

$\sigma^2 = \overline{x^2} - \overline{x}^2$ = 2- $(4/3)^2$ = 2/9

Example-3 Joint Discrete Random Variables

Consider the following discrete joint distribution;

$P(x_i, y_j) =$

	$y_1 = 0$	$y_2 = 1$	$y_3 = 2$
$x_1 = -1$	1/4	1/8	0
$x_2 = 0$	1/16	1/16	1/8
$x_3 = 1$	0	1/8	1/8
$x_4 = 2$	1/16	0	1/16

a) Find the sample space for each RV (x and y) and their probabilities.
b) Find $P(x_i > 1)$, $P(x_i < 1)$, $P(y_j < 0)$, $P(x_i = 1 , y_j = 1)$, $P(x_i > 1 , y_j < 2)$.
c) Is $(x_i = 2$ and $y_j = 1)$ independent or mutual exclusive?
d) Repeat (c) for $(x_i = 1$ and $y_j = 2)$ and $(x_i = 1$ and $y_j = 0)$.

Solution:

a-Using the relation $P(x_i) = \sum_y P(x_i, y_j)$ i.e sum all probabilities for fixed x_i and different y_j (also by summing each row content of above table). Then using the relation $P(y_j) = \sum_x P(x_i, y_j)$ i.e sum all probabilities for fixed y_j and different x_i (also by summing each column content of above table).

X	x_1	x_2	x_3	x_4
x_i	-1	0	1	2
$p(x_i)$	3/8	1/4	1/4	1/8

Y	y_1	y_2	y_3
y_j	0	1	2
$p(y_j)$	3/8	5/16	5/16

One should check the followings as well:

$$\sum_{i=1}^{N} P(x_i) = 1 \qquad \sum_{j=1}^{M} P(y_j) = 1 \qquad \sum_x \sum_y P(x_i, y_j) = 1$$

b-$P(x_i > 1) = P(x_i = 2) = P(x_4) = 1/8$ (from table of p(x))

$P(x_i < 1) = P(x_i = 0) + P(x_i = -1) = P(x_2) + P(x_1) = 1/4 + 3/8 = 5/8$ (from table of p(x))

$P(y_j < 0) = 0$ (why)

$P(x_i = 1 , y_j = 1) = P(x_3 , y_2) = 1/8$ (from table of p(x,y))

$P(x_i > 1 , y_j < 2) = P(x_i = 2 , y_j = 1) + P(x_i = 2 , y_j = 0) = 1/16$ (from table of p(x,y)) c-To see if ($x_i = 2$ and $y_j = 1$) are independent or mutual exclusive, we look at table of p(x,y), it is clear that $P(x_i = 2 \text{ and } y_j = 1) = 0$ thus $x_i = 2$ and $y_j = 1$ are mutually exclusive.

d-For ($x_i = 1$ and $y_j = 2$) from table p(x,y) , $P(x_i = 1 \text{ and } y_j = 2) = 1/8$ then they are not mutual exclusive to see whether they are independent we should find ;

$P(x_i = 1) = 3/8$ and then $P(y_j = 2) = 5/16$ from tables of p(x) and p(y), respectively.

Now since $P(x_i = 1 \text{ and } y_j = 2) \neq P(x_i = 1) \,.\, P(y_j = 2)$, thus they are not independent.

For $P(x_i = 1 \text{ and } y_j = 0) = 0$ from table p(x,y) then they are mutually exclusive.

Q- For Example-3 find the values of $\overline{x}$, $\overline{x^2}$, σ_x^2 , $\overline{y}$, $\overline{y^2}$, σ_y^2