Chapter 4
Random Variables and their Expected Values

4.1 Random Variable

When selecting three items from factory production line, in some cases the quality engineer is interested to know how many devices may be selected. The sample space should be:

$$S = \{ NNN , NND , NDN , NDD, DNN, DND , DDN , DDD\}$$

The engineer tracks the number as D's. which represents the number of defective devices. This number is called the random variable. is denoted by X. its corresponding value is x

Example: Basket contains 4 red balls, 3 black balls, two balls are drawn randomly, count the possibility of getting red ball.

Solution: Call R : Red B : Black , We have the following possibilist.

Sample space	**X**
RR	2
RB	1
BR	1
BB	0

There are two types of random variables

- Discrete.
- Continuous

Discrete random variable is that with countable state such as the number of defective devices. *Continues random variable* takes values on continuous scale such as tracking height and body temperature.

4.2 Discrete Probability Distribution

Here we study the probability distribution of the random variable. The student may be interested in calculating the probability of getting for example 2 Defective devices.

Example: Unbiased coin is tossed twice. The probability of getting a tail is twice the probability of getting head. Calculate the probability distribution of the number of heads.

Solution: given $P\ (\ H\ ) = \frac{1}{2}\ P\ (T)$, we know from probability that *P(H) + P(T) = 1*

Then $P(H) = {}^{1}/_{3}\ , P\ (T) = {}^{2}/_{3}$

Suppose that X is the number of heads then, $S = \{\ HH\ , HT\ , TH\ , TT\ \}$

$$P\ (HH) = P(H) \times P\ (H) = {}^{1}/_{3} \times {}^{1}/_{3} = \frac{1}{9}$$

$$\Rightarrow P\ (\ x\ = 2\ ) = \frac{1}{9}$$

$$P\ (HT) = P(TH) = P(x\ = 1\ ) = P(H)P(T) = {}^{1}/_{3} \times {}^{2}/_{3} = \frac{2}{9}$$

$$P\ (TT) = P(x\ = 0\ ) = P(T) \times P(T) = {}^{2}/_{3} \times {}^{2}/_{3} = \frac{4}{9}$$

$\boldsymbol{X}$	**0**	**1**	**2**	**Total**
$\boldsymbol{P\ (X)}$	${}^{4}/_{9}$	${}^{2}/_{9} + {}^{2}/_{9}$	${}^{1}/_{9}$	1

The set of ordered pairs $\left(\ x, F\ (x)\right)$ is called the probability distribution of the random variable, is called PDF (Probability Distribution Function). PDF has the following properties

- $f\ (x) \geq 0$
- $\sum f\ (x) = 1$
- $P(X\ =\ x) = f(x)$

Another distribution called Commutative distillation function CDF is defined as

$$F(x) = P\ (X \leq x\ ) = \sum\nolimits_{t\ \leq x} f\ (t);\ -\infty < x < \infty$$

Example 2 : Calculate CDF of the above example (1)

$$for\ x < 0\ , f\ (\ x) = 0$$

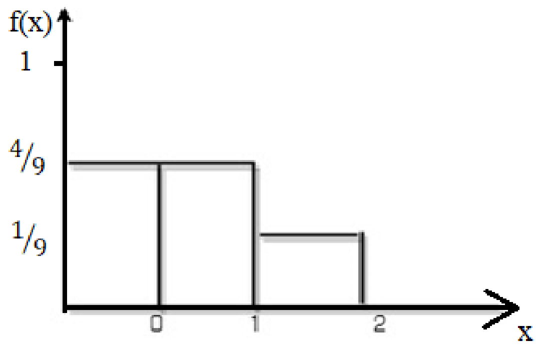
$$for\ 0\ \leq x\ < 1\ , \qquad f\ (x) = \frac{4}{9}$$

$$for\ \ 1 \leq x\ < 2\ ,\ \ f\ (x) = P\ (\ x = 0\ ) + P\ (\ x = 1\ )$$

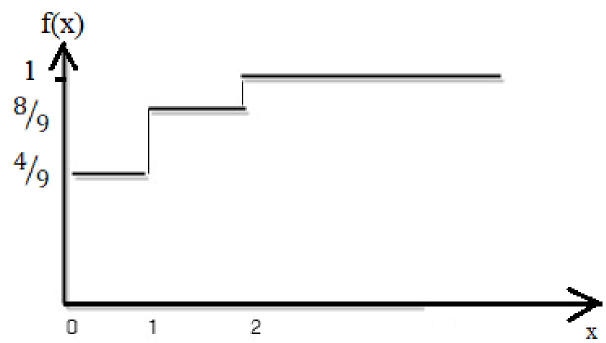
$$= \frac{4}{9} + \frac{4}{9} = \frac{8}{9}$$

$$for\ x \geq 2\ \ \ F\ (x) = P\ (x = 0) + P\ (x = 1) + P\ (x = 2\ )$$

$$= \frac{4}{9} + \frac{4}{9} + \frac{1}{9} = \frac{9}{9} = 1$$



PDF



CDF

4.3 Continuous Probability Distribution

Continuous probability works with continuous variables. For example, if it is required to register patient's body temperature every day. it is found that in a month the degree was between 35 and 38 for 20 days and in 5 days it was less than 35 and for another 5 days it was greater than 38. The resulting measurements are shown in the table

35.183	32.712	38.561	35.509	35.565	35.633	37.587	35.962
35.850	39.291	36.696	35.768	37.293	36.434	36.204	36.559
33.911	34.883	36.780	38.081	38.867	34.332	36.092	38.169
39.542	36.557	38.414	36.146	35.762	34.237		

The measurements in the table can be created with MATLAB command (http://octave-online.net/)

```
>> a=normrnd(36,2,1,30)
```

The frequency table is shown in Figure 2 (Top) with the corresponding histogram (Left), the histogram can be obtained with MATLAB command

```
>> hist(a,6)
```

The curve equation in the figure for PDF must be defined in order to be able to calculate the probability for any period. A question such as, what is the probability that the patients temperature will fall between 36.214 and 37.369 ? The curve equation in this case is called bell curve. The curve follows Normal Distribution (one of the common continuous distribution), its PDF or the curve equation can be defined as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \qquad (1)$$

To draw the curve we don't need the entire data (since we know that it is normally distributed), we need to know only the data's mean and standard deviation. The mean of the patients data is 36.419 and standard deviation is 1.635. The following MATLAB code can be used to plot the curve in Figure 2 (right).

```
>> m=mean(a)
```

```
      m =  36.419
>> sigma=std(a)
      sigma =  1.6357
>> x = (-5 * sigma:0.01:5 * sigma) + m;  %// Plotting range
>> y = exp(- 0.5 * ((x - m) / sigma) .^ 2) / (sigma * sqrt(2 * pi));
>> plot(x,y,'b','linewidth',3)
```

Class center	33.281	34.419	35.558	36.696	37.835	38.973
frequency	1	4	9	7	4	5
Probability	1/30=0.03	4/30=0.13	9/30=0.30	7/30=0.23	4/30=0.13	5/30=0.16

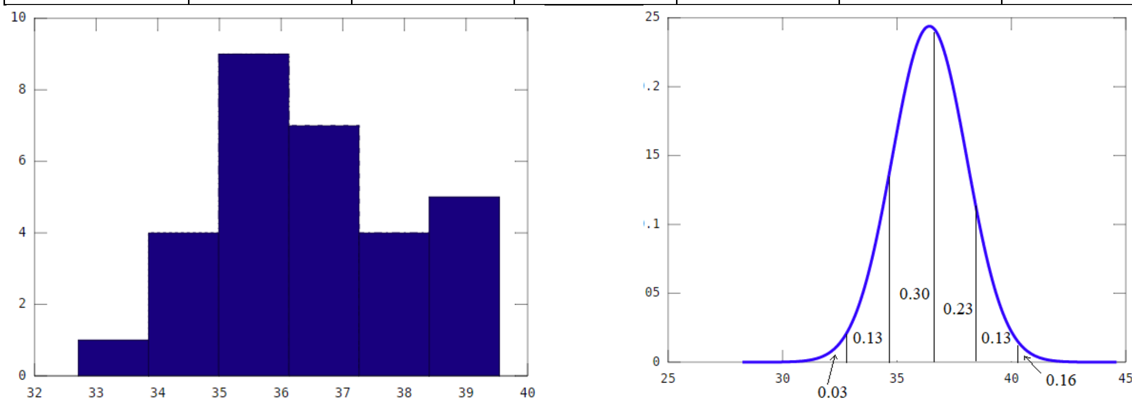


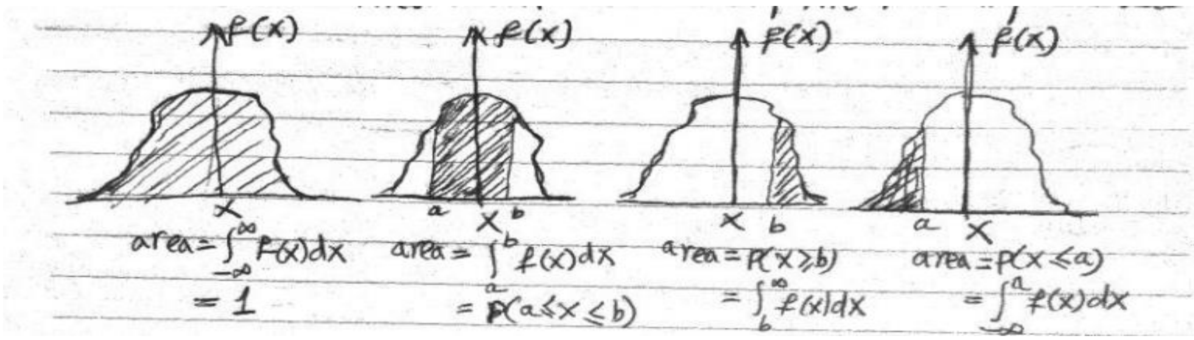
Figure 2: (Top) Frequency table of the patients data, (Left) The histogram of the patients data, (Right) The standard normal curve of the patients data

If for example, you are asked to calculate the probability of getting patient's temperature between a , b . this is an integration problem and must be solved using integration. So area under the curve must be calculated.

$$P\,(a\ <\ x\ <b) = \int_a^b f\,(x)dx$$

It now clear that *f(x)* will be defined by equation (1) only if the data follows Normal Distribution. The probability distribution or probability density function (PDF) for random variable x is defined as a set of real numbers if

1. $f(x) \geq 0 \;\forall\; x \in R\ (R\ is\ real\ numbers\ set)$
2. $\int_{-\infty}^{\infty} f\,(x)dx = 1$
3. $P(\,a < \ x\ < b\,) = \int_a^b f\ (x)dx \;\forall\; a, b\ \in R\,; a\ \leq b$



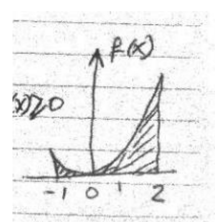
Example : Suppose the error in temperature C° for lab experiment is continues random variable x having the following probability density function .

$$f\ (x) = \begin{cases} 1/3\ x^2\ , & -1\ < x\ < 2 \\ 0, & elsewhere \end{cases}$$

- Verify that $f(x) \geq 0$, And $\int_{\infty}^{\infty} f\ \ (x)\ dx = 1$
- Find $P\ (\ 0\ < x\ \leq 1\ )$

Solution :

- From the drawing , F (x) is quadratic so f (x) > 0



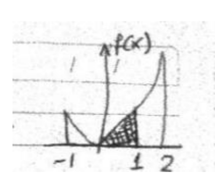
$$\int_{-\infty}^{\infty} f\ \ (x)dx = \int_{-\infty}^{-1} 0\ dx + \int_{-1}^{2} \frac{1}{3}\ x^2\ dx + \int_{2}^{\infty} 0\ dx$$

$$= \int_{-1}^{2} \frac{1}{3}\ x^2\ dx = \frac{1}{9}\ x^3\ |_{-1}^{2}$$

$$= \frac{1}{9}\ (\ 8 - (\ -1\ )) = 1$$

- $P\ (0 < x \leq 1\ ) = \int_0^1 f\ (x)dx = \int_0^1 \frac{1}{3}\ x^2\ dx$

$$= \frac{1}{9}\ x^3\quad |_0^1 = \frac{1}{9}\left(\ 1 - (0)\right) = \frac{1}{9}$$



Another distribution is CDF for continuous random variable, CDF is calculated by

$$F\ (x) = p\ (X\ \leq x\ ) = \int_{-\infty}^{x} f\ (t)dt\ ;\ \ for - \infty\ < x\ < \infty$$

Then

$$p\ (\ a <\ X < b\ ) = P\ (\ X < b\ ) - P\ (X\ < a\ ) = f\ (b) - f(a)$$

Example:

- Find CDF for example 1
- Using CDF, find $P(\ 0 < x\ < 1)$

Solution:

- $f(x) = \int_0^{(1/3)x2} \quad {}^{-1<x<2}_{elsewhere}$

$$\Rightarrow for\ x < -1, f\ (x) = \int_{-\infty}^{x} f\ (t)dt = \int_{-\infty}^{x} 0dt = 0$$

$$\Rightarrow for - 1\ < x\ < 2$$

$$f(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{-1} 0dt + \int_{-1}^{x} \frac{1}{3}\ t^2 dt$$

$$= \tfrac{1}{9}\ t^3\ \big|_{-1}^{x} = \tfrac{1}{9}\ \big(\ x^3 - (\ -1\ )\big) = \tfrac{1}{9}\ (x^3 + 1)$$

$$\Rightarrow for\ x\ > 2$$

$$f(x) = \int_{-\infty}^{x} f\ (t)dt = \int_{-\infty}^{-1} 0dt\ + \int_{-1}^{2} \frac{1}{3}\ t^2\ dt + \int_{2}^{x} 0dt$$

$$= \int_{-1}^{2} \tfrac{1}{3}\ t^2\ dt = \tfrac{1}{9}\ t^3\ \big|_{-1}^{2} = \tfrac{1}{9}\ \big(\ 8 - (-1\ )\big) = 1$$

And then

$$f(x) = \begin{cases} 0 & , \quad x < -1 \\ \dfrac{1}{9}(x^3 + 1\ ) & , \quad -1\ \le x\ < 2 \\ 1 & , \quad x \ge 2 \end{cases}$$

- $P\ (\ 0 < x\ \le 1\ ) = F\ (1) - F\ (0)$

$$= \frac{2}{9} - \frac{1}{9} = \frac{1}{9}$$

4.4 Joint Probability Distribution

We studied single random variable, however there will be situations where we may find it desirable to record the simultaneous outcomes of several random variables. For example, we might measure the amount of precipitate *P* and volume *V* of gas released from a controlled chemical experiment, giving rise to a two-dimensional sample space.

In the discrete case, The function *f*(*x*, *y*) is a *joint probability distribution* or *probability mass function* of the discrete random variables *X* and *Y* if

1. $f(x,y) \ge 0$ for all (x,y),
2. $\sum_x \sum_y f(x,y) = 1$,
3. $P(X = x, Y = y) = f(x,y)$.

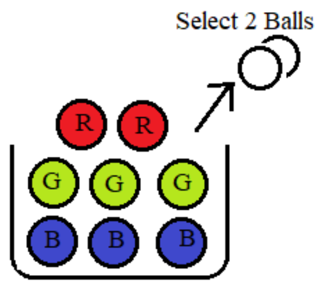
Example: Two ballpoint pens are selected at random from a box that contains 3 blue pens, 2 red pens, and 3 green pens. If *X* is the number of blue pens selected and *Y* is the number of red pens selected, find

(a) the joint probability function $f(x, y)$,
(b) $P[(X, Y) \in A]$, where A is the region $\{(x, y)|x + y \leq 1\}$.
Solution : The possible pairs of values (x, y) are (0, 0), (0, 1), (1, 0), (1, 1), (0, 2), and (2, 0).

(a) Now, $f(0, 1)$, for example, represents the probability that a red and a green pens are selected. The total number of equally likely ways of selecting any 2 pens from the 8 is $\binom{8}{2}= 28$. The number of ways of selecting 1 red from 2 red pens and 1 green from 3 green pens is $\binom{2}{1}\binom{3}{1}= 6$. Hence, $f(0, 1) = 6/28 = 3/14$. Similar calculations yield the probabilities for the other cases, which are presented in Table 3.1. Note that the probabilities sum to 1.

Table : Joint Probability Distribution

	$f(x,y)$	x = 0	1	2	Row Totals
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
Column Totals		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1



(b) The probability that (X, Y) fall in the region A is

$$P[(X, Y) \in A] = P(X + Y \leq 1) = f(0, 0) + f(0, 1) + f(1, 0)$$
$$= \frac{3}{28} + \frac{3}{14} + \frac{9}{28} = \frac{9}{14}.$$

When X and Y are continuous random variables The function $f(x, y)$ is a *joint density function* of the continuous random variables X and Y if

1. $f(x, y) \geq 0$, for all (x, y),
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\ dx\ dy = 1$,
3. $P[(X,Y) \in A] = \int \int_A f(x, y)\ dx\ dy$, for any region A in the xy plane.

Example: A privately owned business operates both a drive-in facility and a walk-in facility. On a randomly selected day, let X and Y, respectively, be the proportions of the time that the drive-in and the walk-in facilities are in use, and suppose that the joint density function of these random variables is

$$f(x,y) = \begin{cases} \frac{2}{5}(2x + 3y), & 0 \leq x \leq 1, 0 \leq y \leq 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Verify condition 2 of the joint density function definition.
(b) Find $P[(X, Y) \in A]$, where $A = \{(x, y) \mid 0 < x < 0.5 , 0.25 < y < 0.5\}$.
Solution:
(a) The integration of $f(x, y)$ over the whole region is

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\ dx\ dy = \int_0^1\int_0^1 \frac{2}{5}(2x+3y)\ dx\ dy$$
$$= \int_0^1 \left(\frac{2x^2}{5} + \frac{6xy}{5}\right)\Big|_{x=0}^{x=1} dy$$
$$= \int_0^1 \left(\frac{2}{5} + \frac{6y}{5}\right) dy = \left(\frac{2y}{5} + \frac{3y^2}{5}\right)\Big|_0^1 = \frac{2}{5} + \frac{3}{5} = 1.$$

(b) To calculate the probability, we use

$$P[(X,Y) \in A] = P\left(0 < X < \frac{1}{2}, \frac{1}{4} < Y < \frac{1}{2}\right)$$
$$= \int_{1/4}^{1/2}\int_0^{1/2} \frac{2}{5}(2x+3y)\ dx\ dy$$
$$= \int_{1/4}^{1/2} \left(\frac{2x^2}{5} + \frac{6xy}{5}\right)\Big|_{x=0}^{x=1/2} dy = \int_{1/4}^{1/2}\left(\frac{1}{10} + \frac{3y}{5}\right) dy$$
$$= \left(\frac{y}{10} + \frac{3y^2}{10}\right)\Big|_{1/4}^{1/2}$$
$$= \frac{1}{10}\left[\left(\frac{1}{2} + \frac{3}{4}\right) - \left(\frac{1}{4} + \frac{3}{16}\right)\right] = \frac{13}{160}.$$

4.5 Mean of Random Variable

Let X be random variable with probability distribution f(x). the mean or expected value of x is denoted by μ_x or E(x) is defined by:

$$E(x) = \mu_x = \begin{cases} \sum_{all\ x} x\, f(x) & if\ x\ is\ diserete \\ \int_{-\infty}^{\infty} x\, f(x) dx\,; & if\ x\ is\ contions \end{cases}$$

If x defined by function g(x) then

$$E(g(x)) = \mu_{g(x)} = \begin{cases} \sum_{all\ x} g(x) f(x); & if\ x\ is\ discrete \\ \int_{-\infty}^{\infty} g(x) f(x)\,; & if\ x\ is\ continous \end{cases}$$

<u>Example</u> 1:

8 computers with 3 defective, for random selection of 2 computers the probability distribution of X where X is the number of defectives is given by .

X	**0**	**1**	**2**	**Total**
$\boldsymbol{f(x)}$	$^{10}/_{28}$	$^{15}/_{28}$	$^{3}/_{28}$	1

Find the expected value of the number of defectives X

<u>Solution</u> **:** $E(x) = \mu_x = \sum_{x=0}^{2} x\ f\ (x)$

$$= (0) \times \frac{10}{28} + (1) \times \frac{15}{28} + (2) \times \frac{3}{28}$$

$$= \frac{15}{28} + \frac{6}{28}$$

$$= \frac{21}{28} = 0.75 \; computers$$

Example2: let X represents life in hours of a lamp, the PDF of X is given by

$$f\,(x) = \begin{cases} \dfrac{20{,}000}{x^3} & ; \;\; x > 100 \\ 0, & \quad elsewhere \end{cases}$$

Find the expected life of the lamps

Solution: $E(x) = \mu_x = \int_{-\infty}^{\infty} x\, f\,(x) dx$

$$= \int_{100}^{\infty} x \;\; \frac{20{,}000}{x^3} \; dx = 20{,}000 \; \int_{100}^{\infty} \frac{1}{x^2} \; dx$$

$$= \; 20{,}000 \; \left[-\frac{1}{x} \Big|_{x=100}^{x=\infty} \right] = -20{,}000 \left[0 - \frac{1}{100} \right] = 200 \; hours$$

We expect the lamp to last for 200 hours .

Example3 : for example 2 suppose X is a function g(x) defined by $g(x) = \frac{1}{x}$, find the expected value .

Solution : $E\left(g(X)\right) = E\left(\frac{1}{x}\right) = \int_{-\infty}^{\infty} g(x) f\,(x) dx = \int_{-\infty}^{\infty} \frac{1}{x} \; f(x) dx$

$$= \int_{100}^{\infty} \frac{1}{x} * \; \frac{20{,}000}{x^3} \; dx = \int_{100}^{\infty} \frac{1}{x^4} * \; 20{,}000 \; dx = \frac{20{,}000}{-3} \; [\frac{1}{x^3} \; |_{x=100}^{x=00}]$$

$$= \frac{-20{,}000}{3} \left[0 - \frac{1}{1{,}000{,}000} \right] = 0.0067 \; (\; hours \;)$$

Example 4 **:** for example 1 , Suppose X is represented by function $g\;(x) = (\; x - 1)^2$

Solution : $E\left(\; g(x)\right) = \mu_{g(x)} = \sum_{x=0}^{2} g\;(x) f(x) = \sum_{x=0}^{2} (x-1)^2 \; f\;(x)$

$$= (0-1)^2 \; f(0) + (1-1)^2 \; f\;(1) + (2-1)^2 \; f(2)$$

$$= (-1)^2 \; \frac{10}{28} + (0)^2 \; \frac{15}{28} + (1)^2 \; \frac{3}{28}$$

$$= \frac{10}{28} + 0 + \frac{3}{28} = \frac{13}{28} \; \# \; computers$$

4.6 Variance of Random Variable

For two different datasets, the mean or expected value maybe equal, but the dispersion of data may differ, so we use another metric called variance of random variable is defined by :

$$Var\ (x) = \ \partial_x^2 = E\ (x-\mu\ )^2\ = \begin{cases} \sum_x (x-\mu)^2\ f\ (x)\ ;\ x\ is\ discrete \\ \int_{-\infty}^{\infty} (x-\mu)^2 f\ (x) dx\ ;\ x\ continous \end{cases}$$

The standard divination $\partial_x = \sqrt{\partial_x^2}$

Example1: Two companies A , B are using random number of cars in a given working day. suppose X is random variable of the number of cars . the distributions of x for company A and B is shown

Company A

x	**1**	**2**	**3**
F(x)	0.3	0.4	0.3

Company B

x	**0**	**1**	**2**	**3**	**4**
F(x)	0.2	0.1	0.3	0.3	0.1

Show that variance of company A data is greater than variance of company B data

Solution: For company A**:**

$$\mu A = E\ (x) = \sum_{x=1}^{3} x\ f\ (x) = (1)(0.3) + (2)(0.4) + (3)(0.3) = 2.0$$

$$\partial_A^2 = Var\ (x) = E\ [(x-\mu A)^2] = \sum_{x=1}^{3} (x-2\ )^2\ f(x)$$

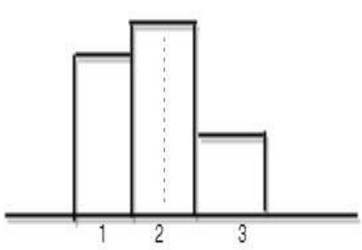
$$= (1-2\ )^2\ (0.3) + (2-2)\ ^2(0.4) + (3-2)^2(0.3) = 0.6$$

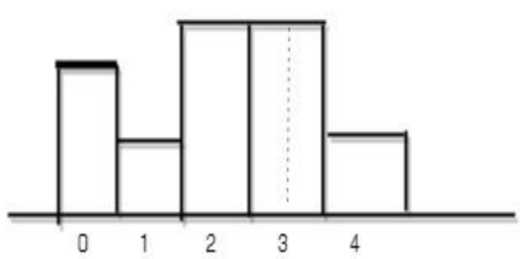
For company B:

$$\mu B = E\ (x) = \sum_{x=0}^{4} x\ f(x) = (0)(0.2) + (1)(0.1) + (2)(0.3) + (3)(0.3) + (4)(0.1) = 2$$

$$\partial_B^2 = Var\ (x) = E\ [(x-\mu B)^2]$$

$$= \sum_{x=0}^{4} (x-2\ )^2\ f(x) = (0-2)^2\ (0.2) + (1-2)^2\ (0.1) + (2-2)^2\ (0.3) + (3 - 2)^2\ (0.3) + (\ 4-2)^2\ (0.1) = \ 1.6$$





Example 2**:** The weekly demand for Pepsi in thousands of litters is continues random variable X has the following PDF.

$$f(x) = \begin{cases} 2(x-1) & ;\ 1< x <2 \\ 0 & ;elsewhere \end{cases}$$

Find the mean and variance of X

<u>Solution:</u>

$\mu = E(x) = \int_{-\infty}^{\infty} x\, f\,(x)dx = \int_{1}^{2} x\,(2(x-1))dx$

$= \int_{1}^{2} x\, 2(x-1)dx = 2\int_{1}^{2} x^2 - x\, dx = 2\left[\frac{x^3}{3} - \frac{x^2}{2}\ \Big|_{x=1}^{x=2}\right] = \frac{5}{3}$

To calculate variance the following compact formula is used

$$\boxed{Var\,(x) = \partial_x^2 = E\,(x^2) - \mu^2}$$

$E(x^2) = \int_{-\infty}^{\infty} x^2\, f\,(x)dx = 2\int_{1}^{2} x^2\,(x-1)dx = 2\left[\frac{x^4}{1} - \frac{x^3}{3}\ \Big|_{x=1}^{x=2}\right] = \frac{17}{6}$

$$Var\,(x) = \partial_x^2 = E\,(x^2) - \mu^2 = \frac{17}{6} - (\frac{5}{3})^2 = \frac{1}{8}$$

4.7 Mean & Var of linear Combination of Random Variables

$if\ x_1, x_2, x_3, \ldots\ldots\ldots. x_n$ are n random variables and a_1, a_2, …. , a_n are constants , then the random variable .

$$X = \sum_{i=1}^{n} a_i\, x_i = a_1\, x_1 + a_2\, x_2 \ldots\ldots\ldots. + a_n\, x_n$$

Is called <u>*linear combination*</u>, it has the following properties.

1. $E\,(a\,X \pm b) \quad = \quad aE\,(X) \pm b$
2. $E\,(X_1 \pm X_2) = E\,(X_1) + E\,(X_2)$
3. $Var\,(aX \pm b) = a^2\, Var\,(X)$
4. $Var\,(aX_1 \pm bX_2) = a^2\, Var\,(X_1) + b^2\, Var\,(X_2)$

Independent X_1, X_2

<u>Example 1</u> **:** Let X be a random variable with the following PDF

$$f\,(x) = \begin{cases} \frac{1}{3}x^2 & ;\ -1 < x < 2 \\ 0 & ,\ elsewhere \end{cases}$$

Find E (4 X + 3)

<u>Solution</u>: $\mu = E(X) = \int_{-\infty}^{\infty} x\, f\,(x)dx = \int_{-1}^{2} x\ \left[\frac{1}{3}\, x^2\right] dx$

$$= \frac{1}{3}\int_{-1}^{2} x^3\, dx$$

$$= \frac{1}{3}\left[\frac{1}{4}\, x^4\ \Big|_{x=-1}^{x=2}\right] = {}^{5}/_{4}$$

E(4X + 3) = 4 E(X) + 3 = 4(5/4) + 3 = 8

<u>Another Solution</u>**:** $E[g\,(x)] = \int_{-\infty}^{\infty} g(x) f\,(x)dx \ \ ; g(x) = 4x + 3$

$E(\,4X + 3) = \int_{-\infty}^{\infty} (\,4\,x + 3\,) f\,(x)dx$

$$= \int_{-1}^{2} (4x+3)\left[\frac{1}{3}x^2\right] dx$$

$$= \int_{-1}^{2} \frac{4}{3}x^3 + x^2\, dx = \left[\frac{x^4}{3} + \frac{x^3}{3}\right]\Big|_{-1}^{2}$$

$$= \frac{16}{3} + \frac{8}{3} - \left(\frac{1}{3} - \frac{1}{3}\right) = \frac{24}{3} - 0 = 8$$

Example 2**:** if X, Y are independent random variables with Var(X) = 2   Var (Y) = 4

$$Z = 3X - 4Y + 8$$

Find Var(Z).

Solution: $Var\,(z) = Var\,(3X - 4Y + 8)$

$$= (3)^2\, Var\,(X) - (4)^2\, Var\,(Y) + 0$$

$$= 9 \times 2 + 16 \times 4 = 82$$

Example 3**:** if X, Y are independent random variables with

$$E\,(X) = 2 \quad Var\,(X) = 4 \quad Var\,(Y) = 1 \quad E(Y) = 7$$

Find

- $E\,(3X + 7)$,$Var\,(3X + 7)$
- $E\,(5X + 2Y - 2)$, $Var\,(5X + 2Y - 2)$

Solution:

- $E\,(3X + 7) = 3E(X) + 7 = 3 \times 2 + 7 = 13$

  $Var(3X + 7) = (3)^2\, Var\,(X) = 9 \times 4 = 36$
- $E(5X + 2Y - 2) = 5\,E\,(X) + 2\,E\,(Y) - 2$

  $$= 5 \times 2 + 2 \times 7 - 2 = 22$$

  $Var\,(5 \times +2Y - 2) = (5)^2\, Var\,(X) + (2)^2 Var\,(Y)$

  $$= 25 \times 4 + 4 \times 1 = 104 \,\#$$

4.8 Mean and Covariance of Two Random Variables

Let X and Y be random variables with *joint probability distribution* $f(x, y)$. The **mean**, or expected value, of the random variable $g(X, Y)$ is

$$\mu_{g(X,Y)} = E[g(X,Y)] = \sum_x \sum_y g(x,y) f(x,y)$$

Dependent

X , Y

if X and Y are discrete, and

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

if X and Y are continuous.

<u>Example</u>: Let X and Y be the random variables with joint probability distribution indicated in The Table. Find the expected value of $g(X, Y) = XY$.

	$f(x,y)$	x = 0	1	2	Row Totals
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
Column Totals		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

***Solution* :** By the Definition we write

$$E(XY) = \sum_{x=0}^{2}\sum_{y=0}^{2} xyf(x,y)$$

$= (0)(0)f(0,0) + (0)(1)f(0, 1)$
$+ (1)(0)f(1, 0) + (1)(1)f(1,1) + (2)(0)f(2, 0)$

$$= f(1,1) = \frac{3}{14}.$$

Let X and Y be random variables with joint probability distribution $f(x, y)$. The **covariance** of X and Y is

$$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = \sum_x \sum_y (x - \mu_X)(y - \mu_y)f(x,y)$$

if X and Y are discrete, and

$$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_y)f(x,y)\, dx\, dy$$

if X and Y are continuous.
The covariance between two random variables is a measure of the nature of the association between the two. The *sign* of the covariance indicates whether the relationship between two dependent random variables is positive or negative. When X and Y are statistically independent, it can be shown that the covariance is zero.

- *<u>Positive sign</u>* : Large values of X result in large values of Y and vice versa
- *<u>Negative sign</u>*: Large values of X result in small values of Y and vice versa
- *<u>Zero</u>*: X, Y are statistically independent.

The alternative and preferred formula for σXY The covariance of two random variables X and Y with means μX and μY , respectively, is given by

$$\sigma_{XY} = E(XY) - \mu_X \mu_Y.$$

<u>Example</u> prev. example describes a situation involving the number of blue refills X and the number of red refills Y. Two refills for a ballpoint pen are selected at random from a certain box, and the following is the joint probability distribution:

	$f(x,y)$	x: 0	1	2	$h(y)$
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
	$g(x)$	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Find the covariance of X and Y .

Solution **:** From Example 4.6, we see that $E(XY) = 3/14$. Now

$$\mu_X = \sum_{x=0}^{2} xg(x) = (0)\left(\frac{5}{14}\right) + (1)\left(\frac{15}{28}\right) + (2)\left(\frac{3}{28}\right) = \frac{3}{4},$$

And

$$\mu_Y = \sum_{y=0}^{2} yh(y) = (0)\left(\frac{15}{28}\right) + (1)\left(\frac{3}{7}\right) + (2)\left(\frac{1}{28}\right) = \frac{1}{2}.$$

Therefore,

$$\sigma_{XY} = E(XY) - \mu_X\mu_Y = \frac{3}{14} - \left(\frac{3}{4}\right)\left(\frac{1}{2}\right) = -\frac{9}{56}.$$

Exercises4

Question1: A shipment of 11 similar microcomputers to a retail outlet contains 4 that are defective. If a school makes a random purchase of 3 of these computers, find

a. the probability distribution of number of defectives
b. The expected value of number of defectives

Question2: For the following distribution of a random variable (5 Marks)

x	*0*	*1*	*2*
f(x)	$\frac{12}{25}$	$\frac{10}{25}$	$\frac{3}{25}$

a) Find the cumulative distribution function (CDF)
b) Find the expected value of the random variable E[x]

Question3: If x and y are independent random variables with: $\mu_x = 3$, $\mu_y = 6$, $\sigma^2_x = 1$, $\sigma^2_y = 4$:

- Find E(3X+7), Var(5X+2y-2).

Question4: The probability density function of X, the lifetime of a certain type of electronic device (measured in hours), is given by f(x) as shown. *Find E(1/X).*

$$F(x) = \begin{cases} \frac{10}{X^2} & ;\ X > 10 \\ 0 & ; \end{cases}$$

Question5: The shelf life, in days, for bottles of a certain prescribed medicine is a random variable having the density function

$$f(x) = \begin{cases} \frac{200}{(x+100)^2}, & x > 0, \\ 0, & elsewhere. \end{cases}$$

Find the probability that a bottle of this medicine will have a shell life of
(a) anywhere from 80 to 120 days.

Question6: The proportion of people who respond to a certain mail-order solicitation is a random variable X having the density function given in

$$f(x) = \begin{cases} \frac{2(x+3)}{7}, & 0 < x < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the expected value of X.

Question7: Complete the following sentences

1. The mean of discrete random variable can be calculated with ……………………….
2. The Variance of discrete random variable can be calculated with ……………………
3. The expected value of discrete random variable is calculated with ……………………
4. The standard deviation of the discrete random variable can be calculated with ……………………….

Question8: If x is a random variable with pdf $f(x) = x \quad , 1 < x < 2$ Find Var (x)

Question9: If X is a random variable with the following distribution

X	1	2	3	4	5
$f(X)$	c	2c	3c	c	3c

1- Find the value of c
2- Find $E(X)$

Multiple Choices Questions:

1) You are given the following probability distribution:

x	0	1	2	3	4
f(x)	0.2	0.1	0.3	0.3	0.1

What is the variance of X?

a. 0.60
b. 2.03
c. 1.59
d. 0.94
e. 0.48

2) random variable is said to be discrete if:

a) its outcomes are countable
b) it can assume any real number within an interval
c) its outcomes are uncountable
d) None of above

3) Which of the following cannot be a probability?

e) a) -1 b) 0 c) 1 d) .05

Answers

Question1: A shipment of 11 similar microcomputers to a retail outlet contains 4 that are defective. If a school makes a random purchase of 3 of these computers, find

a. the probability distribution of number of defectives
b. The expected value of number of defectives

(a) اجمالى $\underline{11} \rightarrow 4D \rightarrow 7ND$

الحل اختيار 3 من 11 بشكل عام يعطى الأشكال الآتية

$$\binom{11}{3} = \frac{11!}{8!\times 3!} = \frac{11\times\cancel{10}^{5}\times\cancel{9}^{3}\times\cancel{8!}}{\cancel{8!}\times\cancel{6}\ \cancel{3!}}$$

$$= 11\times 5\times 3 = 15\times 11 = 165$$

for $\underline{x=0}$ نختار كل المنتجات سليمة من 7 السليم

$$\binom{7}{3}\binom{4}{0} = \frac{7!}{3!\times 4!} = \frac{7\times\cancel{6}\times 5\times\cancel{4!}}{\cancel{6}\times\cancel{4!}} = 35$$

for $\underline{x=1}$ نختار واحد من العاطبة و ~~...~~ الباقى من السليم

$$\binom{7}{2}\binom{4}{1} = \frac{7!}{5!\times 2!}\times 4 = \frac{7\times\cancel{6}^{3}\times\cancel{5!}}{\cancel{5!}\times\cancel{2}}\times 4 = 21\times 4$$

$$= 84$$

for $\underline{x=2}$

$$\binom{7}{1}\binom{4}{2} = 7\times\frac{4!}{2!\times 2!} = \frac{4\times 3!}{4} = 3!\ \cancel{6}$$

$$= 3!\times 7 = 6\times 7$$

$$= 42$$

for $\underline{x=3}$

$$\binom{7}{0}\binom{4}{3} = 1\times\frac{4!}{1!\times 3!} = \frac{4\times\cancel{3!}}{\cancel{3!}} = \underline{4}$$

then distribution for x

x	0	1	2	3
f(x)	$\frac{35}{165}$	$\frac{84}{165}$	$\frac{42}{165}$	$\frac{4}{165}$

(b)

$$E(x) = 0 \times \frac{35}{165} + 1 \times \frac{84}{165} + 2 \times \frac{42}{165} + 3 \times \frac{4}{165}$$

$$= \frac{84}{165} + \frac{84}{165} + \frac{12}{165} = \frac{180}{165}$$

1- If X is a random variable with pdf $f(x) = x \quad 1 \leq x \leq 3$. Find $E(X)$

$f(x) = x$

$$E(x) = \int_1^3 x f(x)\, dx = \int_1^3 x \cdot x\, dx$$

$$= \int_1^3 x^2\, dx$$

$$= \left.\frac{x^3}{3}\right|_1^3 = \frac{(3)^3}{3} - \frac{(1)^3}{3}$$

$$= \frac{3 \times 9}{3} - \frac{1}{3} = \frac{3 \times 9 - 1}{3} = \frac{27-1}{3} = \frac{26}{3}$$