Assignment #1 Solution. **Date: 13th May, 2020**

Statistics and Probability

1. The table below shows the frequency distribution of marks obtained in a test by a group of senior six students in a certain school;

Marks	**10-**	**20-**	**30-**	**40-**	**50-**	**60-**	**70-**	**80 - 90**
frequency	**18**	**34**	**58**	**42**	**24**	**10**	**6**	**8**

(a) Estimate the mean mark.

(b) Draw the cumulative frequency curve. From your graph, estimate

(i) the median mark

(ii) how many would fail if the pass mark is fixed at 40.

(iii) the range of values within which the middle 50% of the insect lies.

Marks	f	**x**	**fx**	**C.f**
10-20	18	15	270	18
20-30	34	25	850	52
30-40	58	35	2030	110
40-50	42	45	1890	152
50-60	24	55	1320	176
60-70	10	65	650	186
70-80	6	75	450	192
80-90	8	85	680	200
	∑f=200		**∑fx=8140**	

(a) Mean, $\bar{x} = \frac{8140}{200} = 40.7$

(b) (i) Median = 38

(ii) 110 students would fail

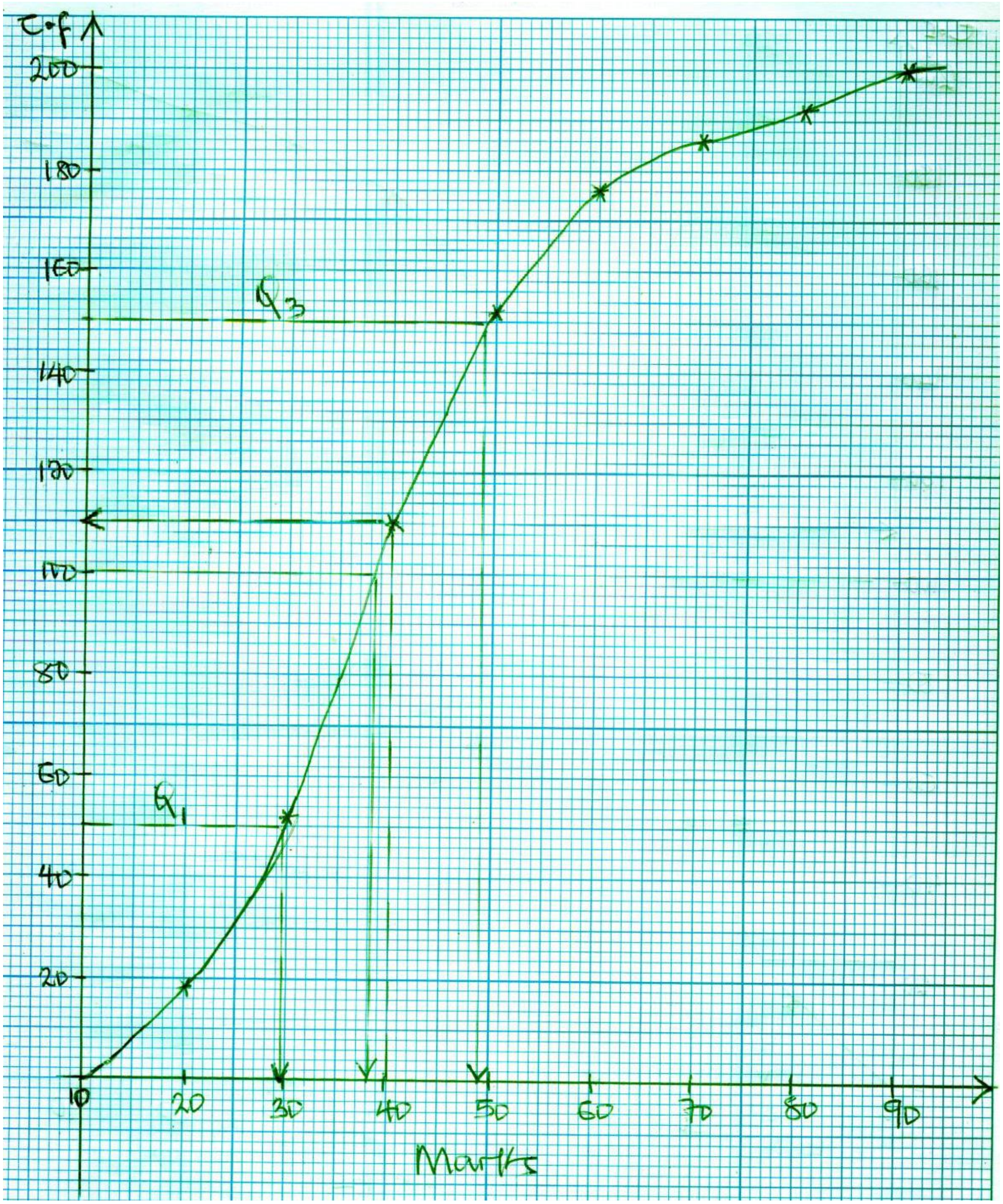
(iii) Middle 50%

P_{75} =75% x200th =150th = 49

P_{25}= 25% x200th =50th = 30

Hence range is (49 – 30)

Prepared by Theode Niyirinda, GHS, 2020. Tel: 0776 286 483 Email: niyirinda@gmail.com



2. A biased tetrahedral die has its faces numbered 1, 2, 3, and 4. If this die is tossed, the probability of the face that it lands on, is inversely proportional to the number on the face. If x is the random variable the number on the face the die lands on, determine;

(a) the probability distribution for x,
(b) Var(3X),
(c) P$[|(x-2)| \leq 1]$

Prepared by Theode Niyirinda, GHS, 2020. Tel: 0776 286 483 Email: niyirinda@gmail.com

(a) $P(X = x) \alpha \frac{1}{x}$ Hence $P(X = x) = \frac{k}{x}$

x	1	2	3	4
P(X=x)	k	$k/2$	$k/3$	$k/4$

$$k + k/2 + k/3 + k/4 = 1$$

$$k = 12/25$$

$$P(X = x) = \begin{cases} \frac{12}{25x}, & x = 1,2,3,4 \\ 0 & elsewhere \end{cases}$$

or

x	1	2	3	4
P(X=x)	$\frac{12}{25}$	$\frac{12}{50}$	$\frac{4}{25}$	$\frac{3}{25}$

(b) $Var(3X) = 3^2 Var(X)$

But $Var(X) = E(X^2) - (E(X))^2$

$$= 1^2\left(\frac{12}{25}\right) + 2^2\left(\frac{12}{50}\right) + 3^2\left(\frac{4}{25}\right) + 4^2\left(\frac{3}{25}\right) - \left[1\left(\frac{12}{25}\right) + 2\left(\frac{12}{50}\right) + 3\left(\frac{4}{25}\right) + 4\left(\frac{3}{25}\right)\right]^2$$

$$= \frac{24}{5} - \left(\frac{48}{25}\right)^2 = \frac{696}{625}$$

Hence Var(3X) = 9x$\frac{696}{625}$ = 10.0224

(c) $P(|(x-2)| \leq 1 = \ P(-1 \leq x - 2 \leq 1)$

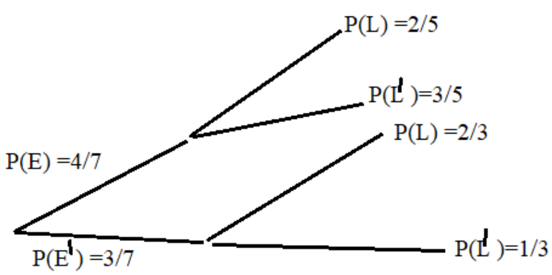
$$= P(1 \leq x \leq 3) = \frac{12}{25} + \frac{12}{50} + \frac{4}{25} = \frac{22}{25} or\ 0.8800$$

3. The probability that the morning chapel at Gayaza High School begins early is $4/7$. If the chapel begins early, the probability that it takes longer is $2/5$. If the chapel begins late, the probability that it takes a shorter time is $1/3$. Find the probability that the chapel

(i) takes a shorter time,

(ii) begins early given that it takes a shorter time.

Let E and L be events chapel begins early and takes longer respectively.



(i) $P(L^1) = P(EnL^1) + P(E^1nL^1)$

$$= \frac{4}{7} x \frac{3}{5} + \frac{3}{7} x \frac{1}{3}$$

$$= \mathbf{\frac{17}{35}}$$

(ii) $P\left(E/_{L^1}\right) = \frac{P(EnL^1)}{P(L^1)} = \frac{\frac{4}{7} x \frac{3}{5}}{\frac{17}{35}}$

$$= \frac{12}{17}$$

Prepared by Theode Niyirinda, GHS, 2020. Tel: 0776 286 483 Email: niyirinda@gmail.com

4. The table below shows the likelihood of where Agnes (A) and Brenda (B) spend their Saturday evening:

	A	***B***
Goes to fellowship	$^1/_2$	$^2/_3$
Visits neighbour	$^1/_3$	$^1/_6$
Stays at home	$^1/_6$	$^1/_6$

a) Find the probability that both go out
b) If we know that they both go out, what is the probability that both go to fellowship?

Let P(G) be probability of both going out i.e fellowship and visiting a neighbor

$$(\boldsymbol{a})\quad \boldsymbol{P(G)} = \frac{1}{2}x\frac{2}{3} + \frac{1}{3}x\frac{1}{6} + \frac{1}{2}x\frac{1}{6} + \frac{2}{3}x\frac{1}{3} = \boldsymbol{\frac{25}{36}}$$

(b) Let F be event of going to the fellowship

$$\boldsymbol{P(^F/_G) = \frac{P(FnG)}{P(G)} = \frac{\frac{1}{3}x\frac{2}{3}}{\frac{25}{36}} = \frac{12}{25}}$$

5. Ten shops in Kampala which attract similar customers are ranked in times of quality of service, size of verandah and price of items. Rank 1 indicates best service, largest verandah and lowest price of commodities. The results, including monthly average sales are given below:

Shop	**Quality of service**	**Size of verandah**	**Price of commodities**	**Sales (kg)**
A	3	3	6	20
B	7	5	10	10
C	4	10	7	31
D	6	7	2	47
E	8	2	4	37
F	2	1	5	38
G	5	8	3	38
H	9	6	8	15
I	10	4	10	21
J	1	9	1	42

a) By calculation, determine whether the price of commodities or the size of the verandah is the more important factor affecting sales.

b) Is there any evidence at 5% level of significance that,

(i) the size of the verandah influences the quality of service?

(ii) a shop with lower priced commodities offer poor quality service (e.g. by employing fewer sales people)?

Let R_Q, R_v, R_P and R_s be ranks for Quality of service, Size of Verandah, Price of Commodities and Sales respectively

Shop	R_Q	R_V	R_P	R_S	D^2_{PS}	D^2_{VS}	D^2_{VQ}	D^2_{PQ}
A	3	3	6	8	4	25	0	9
B	7	5	10	10	0	25	4	9
C	4	10	7	6	1	16	36	9
D	6	7	2	1	1	36	1	16
E	8	2	4	5	1	9	36	16
F	2	1	5	3.5	2.25	6.25	1	9
G	5	8	3	3.5	0.25	20.25	9	4
H	9	6	8	9	1	9	9	1
I	10	4	10	7	9	9	36	0
J	1	9	1	2	1	49	64	0
					$\sum D^2_{PS} = 20.5$	$\sum D^2_{VS} = 204.5$	$\sum D^2_{VQ} = 196$	$\sum D^2_{PQ} = 73$

(a) Correlation coefficient between price of commodities and sales.

$$\rho_{PS} = 1 - \frac{6x20.5}{10(10^2-1)} = 0.8758$$

Correlation coefficient between the size of the verandah and sales.

$$\rho_{VS} = 1 - \frac{6x204.5}{10(10^2-1)} = -0.2394$$

Since there is a highly positive correlation between price of commodities and sales, therefore, it is the price of commodities other than the size of the verandah that affects sale.

(b) (i) Correlation coefficient between size of the verandah and quality of service

$$\rho_{VQ} = 1 - \frac{6x196}{10(10^2-1)} = -0.1879$$

$Since\ |\rho_{VQ}| = 0.1879 < 0.65\ (critical\ or\ table\ value),$ therefore there is no sufficient evidence at 5% level of significance to show that the size of the verandah influences the quality of service.

(ii) Correlation coefficient between price of items and quality of service

$$\rho_{PQ} = 1 - \frac{6x73}{10(10^2 - 1)} = 0.5576$$

Since $|\rho_{PQ}| = 0.5576 < 0.65$, there is no sufficient evidence at 5% level of significance to conclude that a shop with lower priced commodities offer poor quality services

6. The continuous random variable X has a probability function

$$f(x) = \begin{cases} k(x+2); & -1 < x < 0 \\ 2k; & 0 \le x \le 1 \\ k(5-x)/_2; & 1 < x \le 3 \\ 0, & elswhere \end{cases}$$

where k is a constant.

(a) Sketch f(x) and hence find the constant, k
(b) median
(c) $P(\frac{1}{2} < x < 2)$

(a) -1<x<0

f(-1) =k(-1+2) =k

f(0) = k(0+2) = 2k

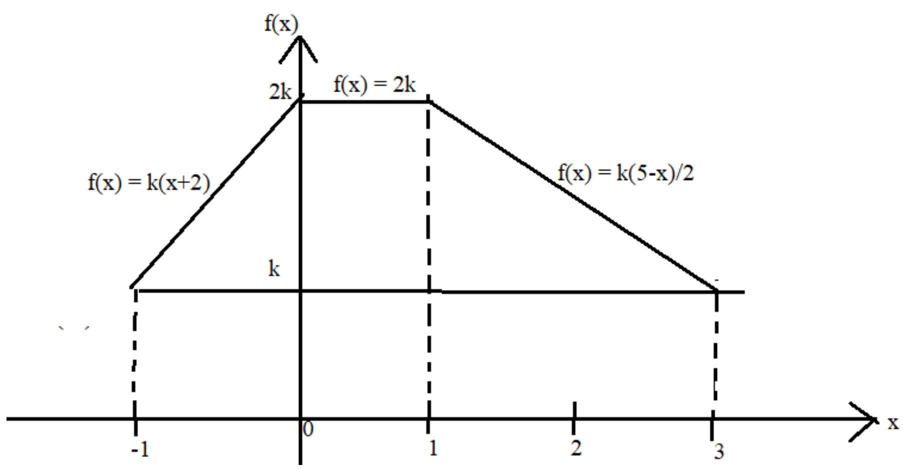
0<x<1

f(0) = f(1) = 2k

1<x<3

f(1) = k(5-1)/2 = 2k

f(3) = k(5 -3)/2 = k



From the graph, total area = 1

$$\frac{1}{2}.k.(4+1) + 4.k = 1 \qquad \therefore k = {}^{2}/_{13}$$

(b) Let m be the median value

$$\int_{-1}^{m} f(x)dx = {}^{1}/_{2},$$

$Testing\ for\ \int f(x)dx > {}^{1}/_{2}$

$$\int_{-1}^{0} \frac{2}{13}(x+2)dx = \frac{2}{13}\left[\frac{x^2}{2} + 2x\right]_{-1}^{0} = \frac{2}{13}\left[0 - \left(\frac{1}{2} - 2\right)\right] = \frac{3}{13} < \frac{1}{2}$$

Prepared by Theode Niyirinda, GHS, 2020. Tel: 0776 286 483 Email: niyirinda@gmail.com

	$\int_{-1}^{0}\frac{2}{13}(x+2)dx + \int_{0}^{1}\frac{4}{13}dx = \frac{3}{13}+\frac{4}{13}=\frac{7}{13}>\frac{1}{2}$ $Therefore\ m\ lies\ in\ the\ range\ of\ 0<x<1$ Hence $\int_{-1}^{0}\frac{2}{13}(x+2)dx + \int_{0}^{m}\frac{4}{13}dx = \frac{1}{2}$ $\frac{3}{13}+\frac{4}{13}(m-0)=\frac{1}{2}$ $\frac{4}{13}m=\frac{7}{26} \quad \therefore m=\frac{7}{8}=0.875$ (c) $P\left(\frac{1}{2}<x<2\right)=\int_{0.5}^{1}\frac{4}{13}dx+\int_{1}^{2}\frac{2}{26}(5-x)dx$ $=\frac{4}{13}[x]_{0.5}^{1}+\frac{2}{26}\left[5x-\frac{x^2}{2}\right]_1^2=\frac{4}{13}(1-0.5)+\frac{2}{26}\left[(10-2)-\left(5-\frac{1}{2}\right)\right]$ $=\frac{11}{26}\ or\ 0.4231$
7.	Events A and B are independent such that P (A) = y, 5P (B) = 5y +1 and 20P($A\cap B$) = 3. (i) Find the value of y. (ii) For this value of y, determine P($A\cup B$) and P(A'/B'). (i) For independent events P(AnB) = P(A) x P(B) $\frac{3}{20}=y.\frac{1}{5}(5y+1) \quad \rightarrow 5y^2+y=\frac{3}{4} \quad \rightarrow 20y^2+4y-3=0 \quad \therefore y=0.3\ or\ y=-0.5$ $Hence\ y = 0.3\ or\ {}^{3}/_{10}$ (ii) $P(AuB)=P(A)+P(B)-P(A).P(B)\quad But\ P(A) = 0.3\ and\ P(B) = \frac{5x0.3+1}{5}=0.5$ = 0.3 +0.5 – 0.3x0.5 = 0.65 $P\left(A^1/_{B^1}\right)=\frac{P(A^1nB^1)}{P(B^1)}$ $=\frac{P(A^1)xP(B^1)}{P(B^1)}$ $=P(A^1)$ $=(1-0.3)$ $=0.7$

Prepared by Theode Niyirinda, GHS, 2020. Tel: 0776 286 483 Email: niyirinda@gmail.com

8. A pharmacist had the following records of unit price and quantities of drugs sold for the years 2010 and 2011.

Drug	Unit price per carton		Quantities per carton	
	2010	**2011**	**2010**	**2011**
Aspirin	80	125	40	45
Panadol	100	90	70	90
Quinine	55	75	8	10
Coartem	90	100	10	10

Taking 2010 as the base year, calculate the;

(i) price relatives for each drug in 2011.

(ii) simple aggregate price index number for 2011.

(iii) weighted price index and comment on it.

(iv) weighted aggregate price index and comment on it.

Drug	P_B	P_c	w	$\frac{P_c}{P_B}$	$\left(\frac{P_c}{P_B}\right)w$	P_Bw	P_cw
Aspirin	80	125	45	1.5625	70.3125	3600	5625
Panadol	100	90	90	0.9	81	9000	8100
Quinine	55	75	10	1.3636	13.636	550	750
Coartem	90	100	10	1.1111	11.1111	900	1000
	$\sum P_0=325$	$\sum P_c=390$	$\sum w=155$		$\sum\left(\frac{P_c}{P_B}\right)w = 176.0596$	$\sum P_Bw = 14050$	$\sum P_cw = 15475$

(i) Refer to the table; Price relative = $\frac{P_c}{P_B}$ or $\frac{P_c}{P_B}x100$

(ii) Simple aggregate price index = $\frac{\sum P_c}{\sum P_B}x100 = \frac{390}{325}x100 = 120$

(iii) Weighted price Index = $\frac{\sum\left(\frac{P_c}{P_B}\right)w}{\sum w}x100 = \frac{176.0596}{155}x100 =$

= 113.5868

Comment: Hence the prices in 2011 increased by 13.5868%

(iv) Weighted aggregate Price Index = $\frac{\sum P_cw}{\sum P_Bw}x100 = \frac{15475}{14050}x100 = 110.1423$

Hence the price in 2011 increased by 10.1423%

Prepared by Theode Niyirinda, GHS, 2020. Tel: 0776 286 483 Email: niyirinda@gmail.com

9.	A random variable X has the probability density function $$f(x)=\begin{cases}\frac{1}{b-a} & a\le x\le b\\ 0 & elsewhere\end{cases}$$ a) Show that the variance of x is $\frac{(b-a)^2}{12}$. b) If the expected value and variance of the distribution is 1 and $\frac{3}{4}$ respectively, find; (i) $P(x<0)$ (ii) value of a given that $P(X > a + \sigma) = \frac{1}{4}$. Where σ is the standard deviation of X. (a) $Var(x) = E(x^2)- (E(x))^2$ But $E(x) = \int_a^b x\left(\frac{1}{b-a}\right)dx = \left(\frac{1}{b-a}\right)\left[\frac{x^2}{2}\right]_a^b = \frac{1}{(b-a)}\left(\frac{b^2-a^2}{2}\right) = \frac{(b-a)(b+a)}{2(b-a)} = \frac{(b+a)}{2}$ $E(x^2) = \int_a^b x^2\left(\frac{1}{b-a}\right)dx = \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b = \frac{1}{b-a}\left(\frac{b^3-a^3}{3}\right)$ $= \frac{1}{b-a}\left(\frac{(b-a)(b^2+ab+a^2)}{3}\right)$ $= \frac{(b^2+ab+a^2)}{3}$ Hence Var(x) $= \frac{(b^2+ab+a^2)}{3} - \left[\frac{(b+a)}{2}\right]^2$ $= \frac{(b^2+ab+a^2)}{3} - \frac{(b^2+2ab+a^2)}{4}$ $= \frac{4(b^2+ab+a^2)-3(b^2+2ab+a^2)}{12}$ $= \frac{4b^2+4ab+4a^2-3b^2-6ab-3a^2}{12}$ $= \frac{b^2-2ab+a^2}{12}$ $= \frac{(b-a)^2}{12}$ **(b)** $\boldsymbol{E(x)} = \frac{(b+a)}{2} = 1 \quad \rightarrow (b+a) = 2 \ldots\ldots\ldots.(1)$ $\boldsymbol{Var(x)} = \frac{(b-a)^2}{12} = \frac{3}{4} \quad \rightarrow (b-a)^2 = 9 \quad \rightarrow (b-a) = 3 \ldots\ldots\ldots.(2)$ **Solving (1) and (2) simultaneously gives** $\boldsymbol{a = -0.5,\ b = 2.5}$

Prepared by Theode Niyirinda, GHS, 2020. Tel: 0776 286 483 Email: niyirinda@gmail.com

(i) $\boldsymbol{P(x<0)=\int_{-0.5}^{0}\frac{1}{3}dx=\left[\frac{x}{3}\right]_{-0.5}^{0}-\frac{(0--0.5)}{3}=\frac{1}{6}}$

(ii) $P(X > a + \sigma) = \frac{1}{4}$ $\quad But\ \sigma = \sqrt{Var(x)} = \sqrt{\frac{3}{4}} = \frac{\sqrt{3}}{2} = 0.8660$

$$\int_{\left(a+\frac{\sqrt{3}}{2}\right)}^{2.5}\frac{1}{3}dx=\left[\frac{x}{3}\right]_{\left(a+\frac{\sqrt{3}}{2}\right)}^{2.5}=\frac{2.5-\left(a+\frac{\sqrt{3}}{2}\right)}{3}=\frac{1}{4}$$

$$2.5-\left(a+\frac{\sqrt{3}}{2}\right)=\frac{3}{4}$$

$$a+\frac{\sqrt{3}}{2}=2.5-\frac{3}{4}$$

$$a=1.75-\frac{\sqrt{3}}{2}$$

$a = 0.5253$

10. The number of days the machine breaks down in a month follows a discrete random variable X with a pdf, f(x) given by;

$$f(x)=\begin{cases}k\left(\frac{1}{4}\right)^{x}, x=0,1,2,\dots\dots.. \\ 0 \qquad else\ where\end{cases}$$

Find the i) value of the constant, k

ii) probability that the machine breaks down not more than 2 times in a month.

(i) $k\left(\frac{1}{4}\right)^{0}+k\left(\frac{1}{4}\right)^{1}+k\left(\frac{1}{4}\right)^{2}+\cdots\dots\dots\dots..=1$

Sum to infinity $S_{\infty}=\frac{a}{1-r}$ where $S_{\infty}=1, a=k\ and\ r=\frac{1}{4}$

$$1=\frac{k}{1-{}^{1}/_{4}} \qquad \therefore k={}^{3}/_{4}$$

(ii) $P(x\le 2)=P(X=0)+P(X=1)+P(X=2)$

$$=\frac{3}{4}\left[\left(\frac{1}{4}\right)^{0}+\left(\frac{1}{4}\right)^{1}+\left(\frac{1}{4}\right)^{2}\right]$$

$$=\frac{63}{64}\ or\ 0.9844$$

Prepared by Theode Niyirinda, GHS, 2020. Tel: 0776 286 483 Email: niyirinda@gmail.com