	ADITYA KUMAR	20211204

S.No.		EXPERIMENT	Sign
1		To compute Survivalship prob, density fn & hazard fn & plot their graph	
2.	A	To compute Survivalship prob density fn & hazard fn & plot their graph	
	B.	To determine survivalship fn, hazard fn mean, median, variance & coefficient of different parameter	
3.		To plot compute Censored time using Kaplan Meir product limit method for survival time.	
4.		To compute the estimate of the survival fn & standard error of estimated survival function.	
5.		To compute hazard fn.	
6.		To compute the crude probability of death & partially probability of death.	
7.		To estimate mean, survival rate time, SE of mean survival rate time	
8.		To compare estimated survival time & variance of estimated survival time	
9.		To compute crude probability of death.	
10.		To compute type I & type II censoring.	
11.		To determine gene frequency.	
12.		To compute probability in gamets.	

PRACTICAL - 01

Aim:→ To compute the survivalship of probability density fn & hazard fn & plot their graph

Problem:→
The following table is life time of the total popln of 1000 living birds in the US from 1959 to 1961. Compute & plot the estimated survivalship fn, pdf & hazard fn.

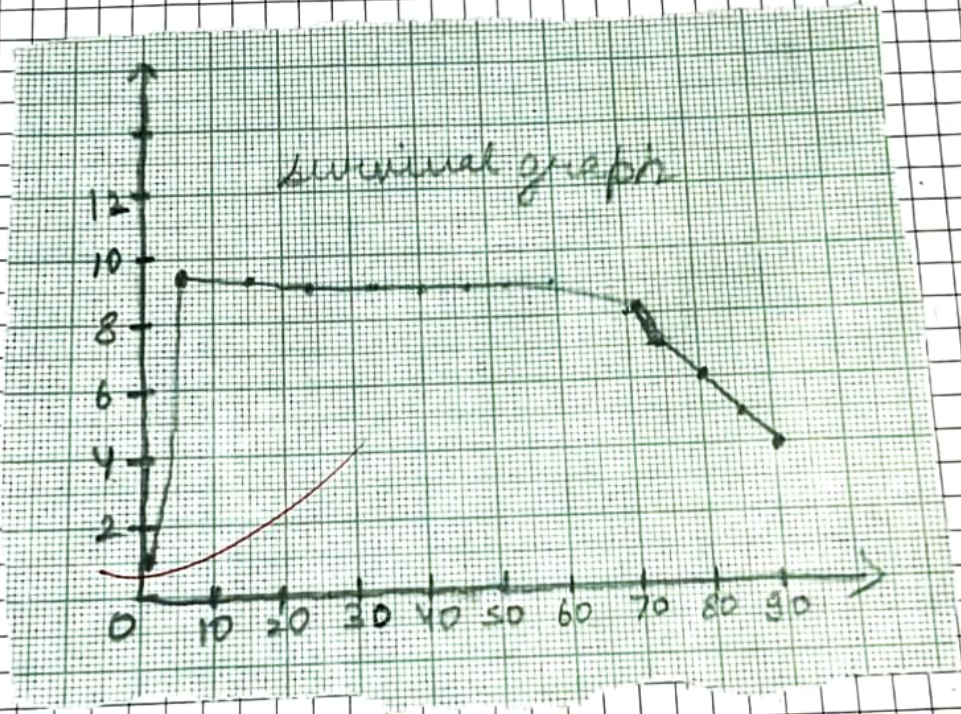
Age	No. of living at begining	No. of dying in interval	Age	No. of living at begining	No. of dying in interval
~~0-1~~					
0-1	10000	2598	45-50	91373	2612
1-5	9~~74~~407	409	50-55	88756	405
5-10	96998	233	55-60	84776	564
10-15	96587	214	60-65	84711	792
15-20	96111	440	65-70	79067	1029
20-25	95517	594	70-75	71197	126
25-30	94905	612	75-80	60857	145
30-35	94144	761	80-85	46170	15
35-40	93064	1080	85≥	33516	18
40-45	96765	1686			

Theory And Formula:→

1) Survival fn:→

$S(t)$ = P [an individual survives longer than t]
= 1 − P(an individual fails before t).

$$\hat{S}(t) = \frac{\text{No. of individual survival longer } t}{\text{Total no. of individuals}}$$


survival graph
12
10
8
6
4
2
0
10
20
30
40
50
60
70
80
90

2) Death / Probability density f^n :→

$$f(t) = \lim_{\Delta t \to 0} \frac{P(\text{an individual lies in time interval } (t, t+\Delta t))}{\Delta t}$$

$$\hat{f}(t) = \frac{\text{No. of individual dying in the interval beg. at time } t)}{(\text{Total no. of individual}) \times (\text{interval width})}$$

3) Hazard function.

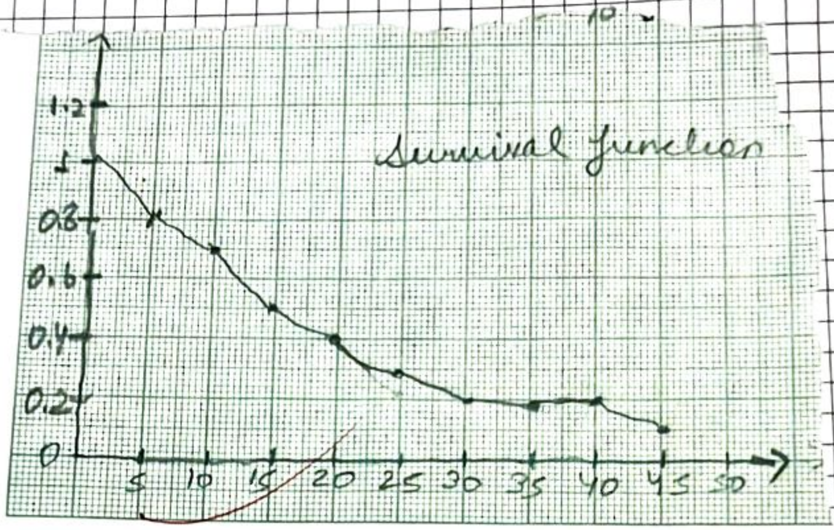
$$h(t) = \lim_{\Delta t \to 0} \frac{P(\text{an individual fails in interval } (t, t+\Delta t) \text{ given that the individual was survived to } t)}{\Delta t}$$

from $f(t)$ & $S(t)$.

we get, $h(t) = \dfrac{f(t)}{S(t)}$.

CALCULATIONS :→

Age	No. of living life at begin.	No. of dying in the interval	S(t) Survival f^n	f(t) Density f^n	h(t) hazard f^n
0-1	10000	2593	1	0.2593	0.25
1-5	97407	409	9.7407	0.0102	0.00
5-10	96998	233	9.6998	0.0046	0.004
10-15	96765	214	9.6765	0.00428	0.000
15-20	96551	440	9.6551	0.0088	0.0
20-25	96111	594	9.6111	0.01188	0.00
25-30	95577	612	9.5577	0.01224	0.00
30-35	94905	761	9.4095	0.01522	0.00
35-40	94148	1080	9.4144	0.02160	0.0e
40-45	93064	1686	9.3064	0.05244	0.00
45-50	91373	2622	9.1373	0.08009	0.00
50-55	88756	4045	8.8756	0.11288	0.00
55-60	84711	5744	8.4956	0.1584	0.0
60-65	79067	7920	7.9061	0.2058	0.0
65-70	71147	10290	7.1147	0.25374	0.0
70-75	60857	12687	6.0857	0.29188	0.3
75-80	46170	14594	4.617	0.1584	0.0
80-85	46176	15034	3.3576	0.29188	0.
82	353578	18590		0.30068	



Survival function
1.2
1
0.8
0.6
0.4
0.2
0
5
10
15
20
25
30
35
40
45
50

PRACTICAL-2A

Aim :→ To compute Survivalship probability density function & hazard function & plot their graph.

Problem :→

Survival data of 40 Myeloma Patients is →

Survival time (t)	No. of patients Survival of beginning	No. of patients dying in interval.
0-5	40	5
5-10	35	7
10-15	28	6
15-20	22	4
20-25	18	5
25-30	13	4
30-35	9	4.
35-40	5	0
40-45	5	2
45-50	3	1
≥ 50	2	2

Theory & Formula :→

1) Survival Function

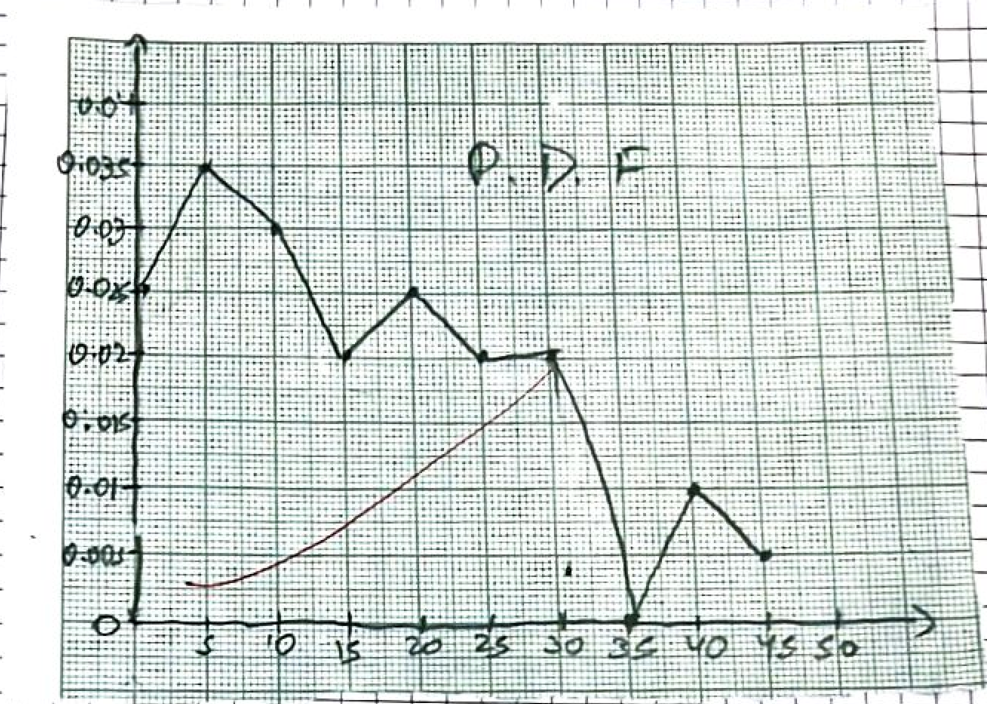
$S(t)$ = P(an individual survives longer than t)

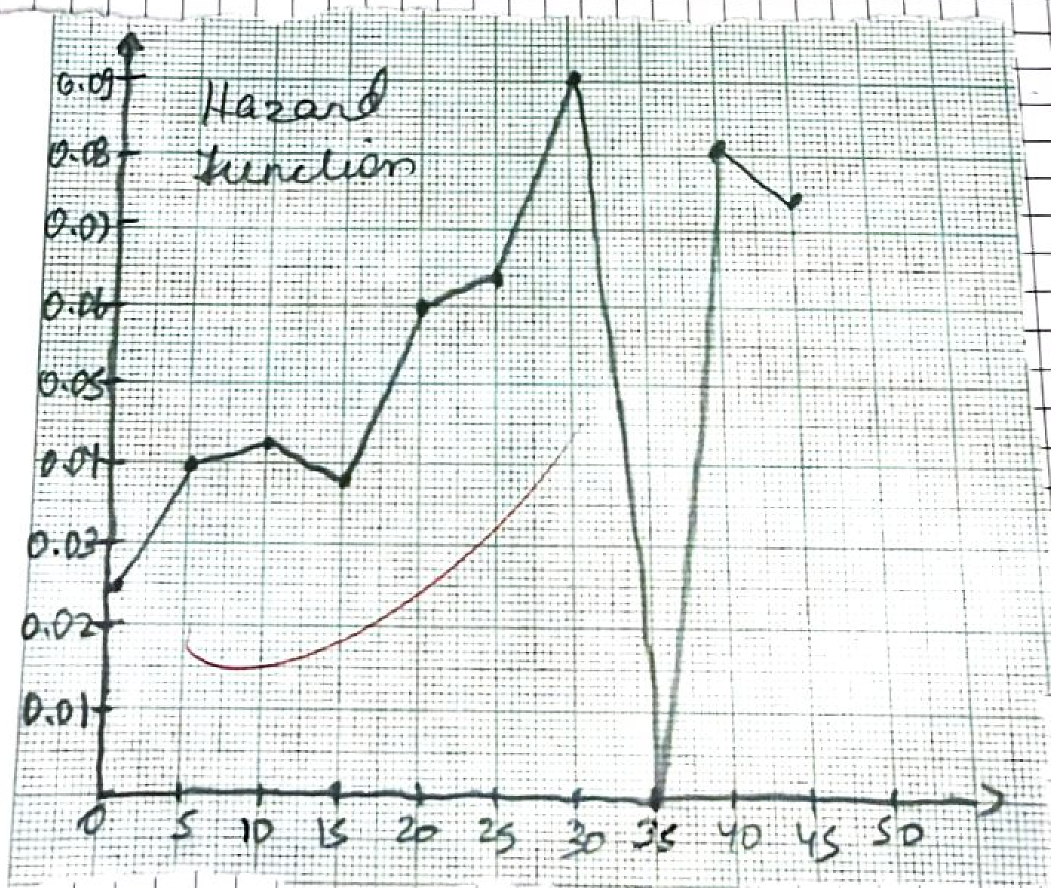
= 1 - P(an individual fails before t).

$$\hat{S}(t) = \frac{\text{No. of individuals surviving longer than } t}{\text{Total no. of individuals}}$$

2) Death / Probability density function

$$f(t) = \lim_{\Delta t \to 0} \frac{P(\text{an individual dies in time interval } t, t+\Delta t)}{\Delta t}$$


P.D.F
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
5
10
15
20
25
30
35
40
45
50


Hazard function
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
5
10
15
20
25
30
35
40
45
50

$$f(\hat{t}) = \frac{\text{No. of individual dying in the interval beg. at time } t}{(\text{Total no. of individual}) \times (\text{interval width})}$$

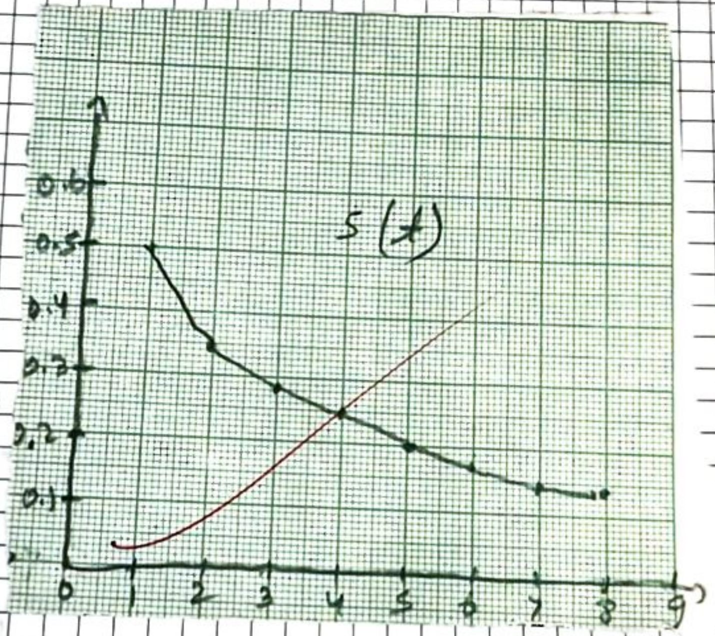
3) Hazard function

$$h(t) = \lim_{\Delta t \to 0} \frac{P\,(\text{an individual fails in the interval } (t, t+\Delta t) \text{ given that the individual survived to } t.}{\Delta t}.$$

$$h(t) = \frac{f(t)}{S(t)}.$$

CALCULATIONS :→

Survivals time t	No. of patients surviving at beginning of interval	No. of patients dying at interval	S(t)	f(t)	h(t)
0-5	40	5	1	0.025	0.0
5-10	35	7	0.875	0.035	0.0
10-15	28	6	0.7	0.03	0.0
15-20	22	4	0.55	0.02	0.0
20-25	18	5	0.45	0.025	0.0
25-30	13	4	0.325	0.02	0.0
30-35	9	4	0.225	0.02	0.
35-40	5	0	0.125	0	0
40-45	5	2	0.125	0.01	0
45-50	3	1	0.075	0.005	0.
≥ 50	2	2	0.05		



S (t)
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
6
7
8
9

Practical No - 3

Aim - To compute the estimate of survival fx^n & standard error of estimated survival fx^n.

Experiment - Suppose that 22 patients surviving time with solid tumour were seen as following -

S. No.	Surviving time	S. No.	Surviving time
1	3.5	3	18
12	5	13	20
7	6.5	2	+8.4
8	6.5	4	+9
5	7	21	+8
6	7	14	+9.5
17	8	15	+11
22	10	18	+12.5
9	12	16	+13
10	15	19	+15
11	16	20	+17

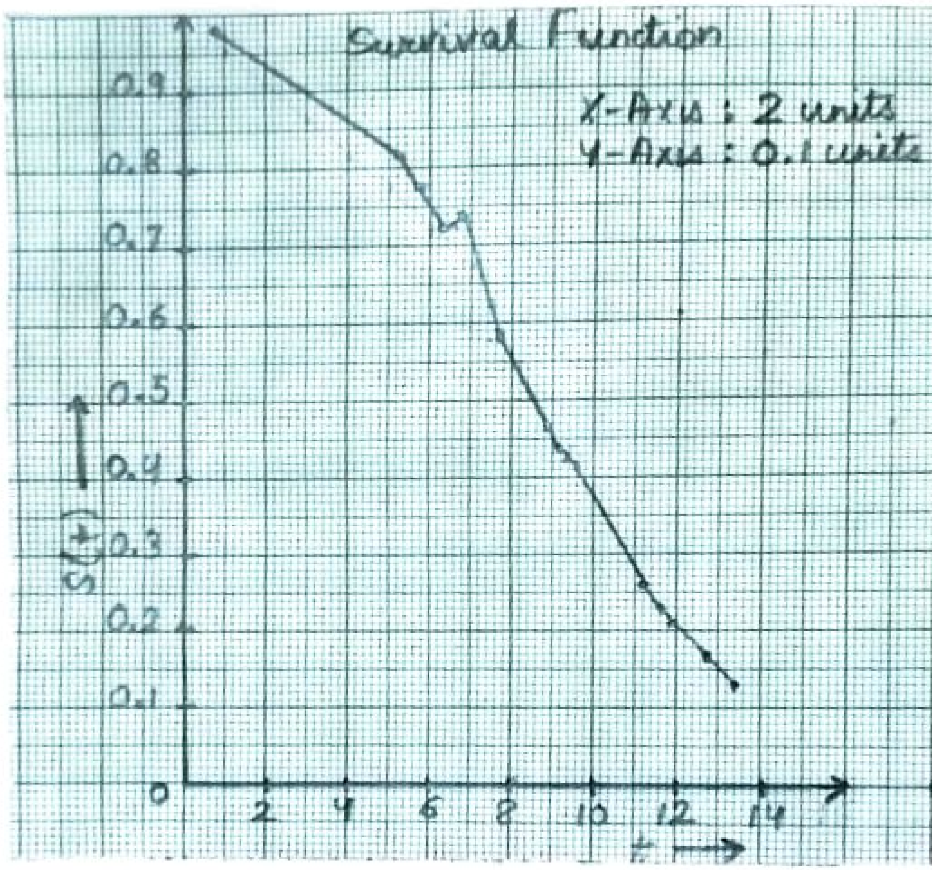
Write the procedure to estimate the survival fx^n & standard error of the estimated survival fx^n. Also, plot the estimated death density fx^n.

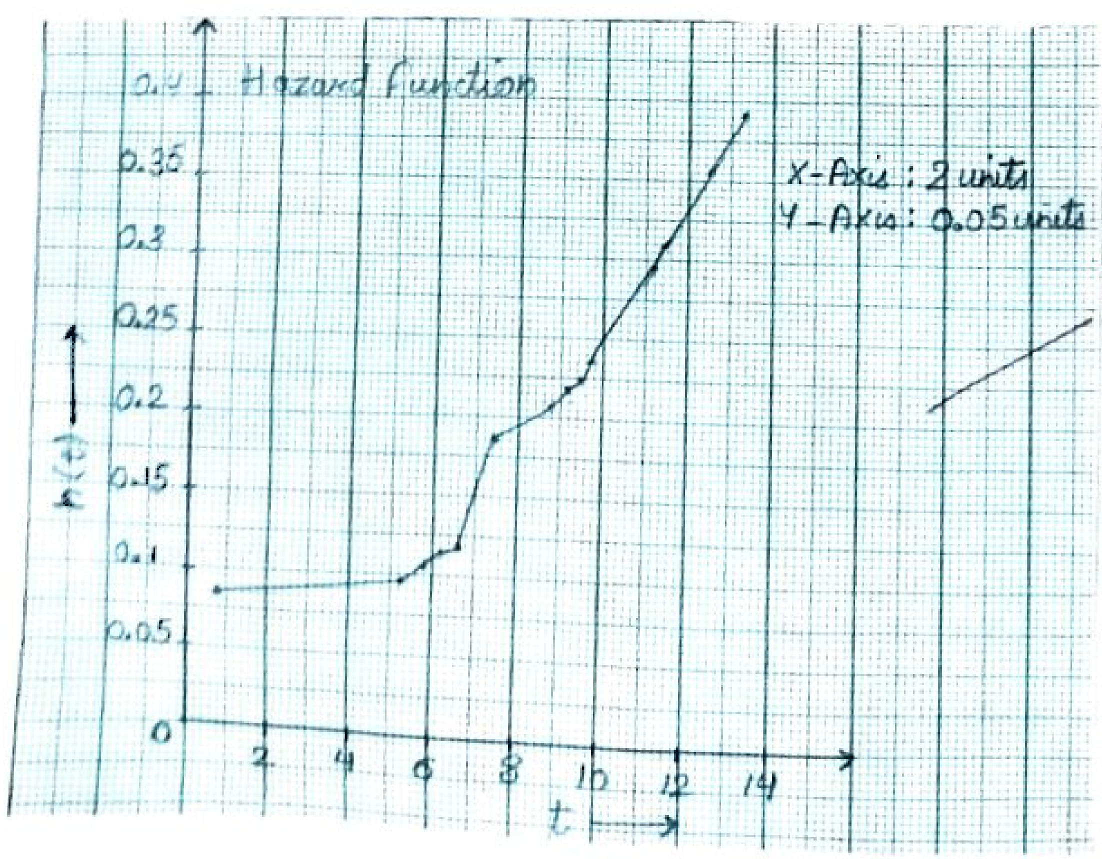
Theory and Formula-

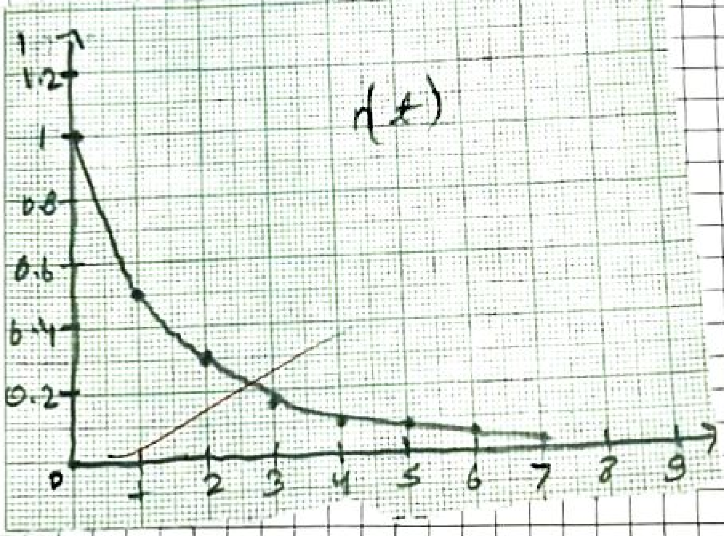
Kaplan Meier Product Limit method -

This method is an extension of lifetime method in which ungroup survival time are given & is based on type-I censored sample.

CLASSTIME


Survival Function
X-Axis : 2 units
Y-Axis : 0.1 units
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2
4
6
8
10
12
14
S(t)
t


Hazard function
X-Axis : 2 units
Y-Axis : 0.05 units
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
2
4
6
8
10
12
14
h(t)
t



PRACTICAL - [2B]

Aim:→ To determine survival function, hazard function mean, median, variance & cofficient of variation of weibull & Expotential distribution of different parameters.

PROBLEM :→

Find mean, variance, cofficient of variation, survival function & hazard function of

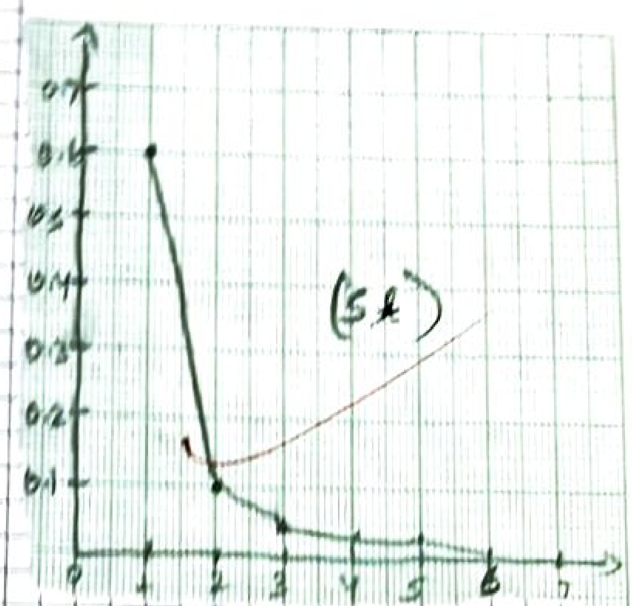
(I) weibull distribution
a) $\lambda = 1$, $\gamma = 0.5$
b) $\lambda = 0.5$, $\gamma = 2$

(II) Exponential Distribution :
$\lambda = 0.65$

Theory & Formula:→

Weibull Distribution:→

Characterised by 2 parameters γ & λ


0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
6
7

$\gamma \to$ determines shape.

$\lambda \to$ determine scale.

$$f(t) = \lambda\gamma(\lambda t)^{\gamma-1} e^{-(\lambda t)^{\gamma}} \; ; \; t \geq 0, \gamma, \lambda > 0$$

$$F(t) = 1 - e^{-(\lambda t)^{\gamma}}$$

$$S(t) = e^{-(\lambda t)^{\gamma}}$$

$$h(t) = \lambda\gamma(\lambda t)^{\gamma-1}$$

$$E(t) = \frac{1}{\lambda}\Gamma\left(\frac{1}{\gamma}+1\right), \quad Var(t) = \frac{1}{\lambda^2}\left[\Gamma\left(\frac{2}{\gamma}+1\right) - \left(\Gamma\left(\frac{1}{\gamma}+1\right)\right)^2\right]$$

Exponential Distribution :-

In exp distribution hazard rate λ is constant where

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & ; t \geq 0, \lambda > 0 \\ 0 & ; t < 0 \end{cases}$$

CDF

$$F(t) = 1 - e^{-\lambda t}$$
$$S(t) = e^{-\lambda t}$$
$$h(t) = \lambda$$
$$E(t) = \frac{1}{\lambda}, \quad Var(t) = \frac{1}{\lambda^2}.$$

Large λ indicates high risk & short survival; small λ indicates low risk & long survival.

CALCULATION :→

$\lambda = 1$

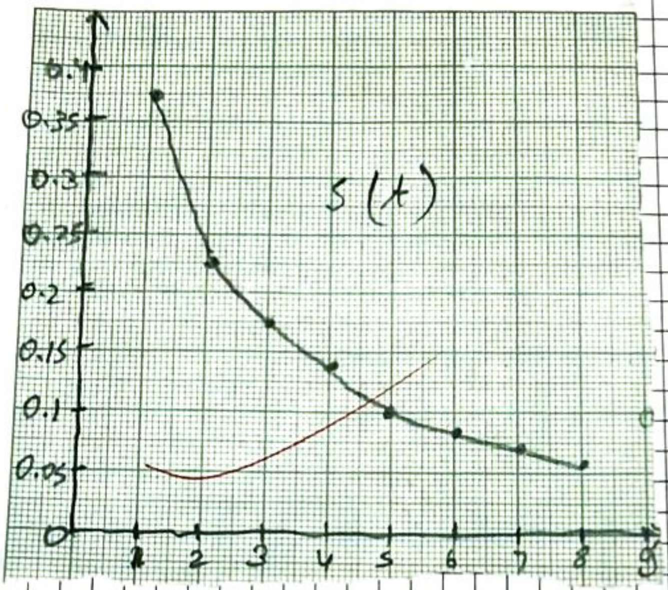
$\gamma = 0.5$

Mean = 2

Variance = 20

Cofficient of variation = 223.60

t	S(t)	h(t)
1	0.3678	0.5
2	0.2431	0.3
3	0.1769	0.2
4	0.1353	0.2
5	0.1068	0.22
6	0.0863	0.2
7	0.0709	0.1
8	0.059106	0.1



s(t)
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
1
2
3
4
5
6
7
8
9

$\lambda = 0.5$

$\gamma = 2$

Mean = 1.772454

Variation = 0.858407

Coeff. of variation = 52.27232

t	S(t)	h(t)
1	0.606531	0.5
2	0.135335	1
3	0.011109	1.5
4	0.00335	2
5	$3.73e^{-06}$	2.5
6	$1.52e^{-08}$	3
7	$2.29e^{-11}$	3.5
8	$1.2e^{-14}$	4

$\lambda = 0.65$

Mean = 1.538462

Coefficient of variation = 100

t	S(t)	h(t)
0	1	0.65
1	0.522046	0.65
2	0.272532	0.65
3	0.142274	0.65
4	0.74274	0.65
5	0.038774	0.65
6	0.020242	0.65
7	0.010567	0.65
8	0.005517	0.65

RESULT :→

PRACTICAL - 3

Aim: To compute censored time using kalplan Meier product limit method for survival time.

Problem:

Person with acute lung problem is given below:

t_i	t_i	t_i	t_i	t_i
2	20	40	60	110
9	25	45	62	110
10	30	45	69	130
19	35	48	89	145
19	40	50	90	160

Theory & Formula:

This method is an extension of life table method in which ungrouped survival ties are given & is based on type I - progressively censored samp

$$p_i = P(T > t_i \mid T > t_{i-1})$$

where, T = r.v presenting survival time of individual

p_i = probability [of surviving in i^{th} sub interv / survival at the begining of sub-int

$$\hat{S}(t_k) = \text{Prob. of surviving atleast time } k.$$

$$= P(T > t_k)$$

$$= \prod_{i=1}^{k} p_i$$

$$p_i = \frac{n-r}{n-r+1}, \quad \hat{S}(t_k) = \prod_{i=1}^{k} \frac{n-r}{n-r+1}$$

PRACTICAL - 3

Aim:→ To compute censored time using kalplan Meier product limit method for survival time.

Problem :→

Person with acute lung problem is given below :→

t_i	t_i	t_i	t_i	t_i
2	20	40	60	110
9	25	45	62	110
10	30	45	69	130
19	35	48	89	145
19	40	50	90	160

Theory & Formula :→

This method is an extension of life table method in which ungrouped survival ties are given & is based on type I - progressively censored sample

$$p_i = P(T > t_i \mid T > t_{i-1})$$

where, T = r.v presenting survival time of individual

p_i = probability [of surviving in i^{th} sub interval / survival at the begining of sub-interval]

$$\hat{S}(t_k) = \text{Prob. of surviving atleast time } k.$$

$$= P(T > t_k)$$

$$= \prod_{i=1}^{k} p_i$$

$$p_i = \frac{n-r}{n-r+1}, \quad \hat{S}(t_k) = \prod_{i=1}^{k} \frac{n-r}{n-r+1}$$

$$Var(\hat{s}(t_k)) = [\hat{s}(t_k)]^2 \sum \frac{1}{(n-r)(n-r+1)}$$

$$S.E[\hat{s}(t_k)] = \sqrt{Var(\hat{s}(t_k))}$$

CALCULATIONS:→

t_i	i	r	p_i	$\hat{s}(t_i)$	$Var(\hat{s}(t_i))$	$SE(\hat{s}(t_i))$
2	1	1	0.96	0.96	0.001536	0.03919
+ 9	2	–		–	–	–
+ 10	3	–		–	–	
19	4	4	0.9545	0.9164	0.002926	0.05691
19	5	5	0.9573	0.8727	0.004035	0.063522
+ 20	6	–	–	–	–	–
25	7	7	0.9474	0.8264	0.004988	0.070624 / 0.07568
30	8	8	0.9445	0.7809	0.005664	0.07803
35	9	9	0.9412	0.7349	0.00609	0.079318
40	10	10	0.9375	0.6889	0.006291	
40	11	–	—	–	—	—
45	12	12	0.9286	0.6396	0.06345	0.079657
45	13	13	0.9231	0.5905	0.006186	0.07865
48	14	14	0.9167	0.5413	0.005849	0.07647
+ 50	15	–	–	–	–	–
60	16	16	0.9	0.98872	0.005364	0.07323
62	17	17	0.8889	0.93308	0.009727	0.0682
69	18	18	0.875	0.3918	0.003997	0.0631
89	19	19	0.8751	0.2706	0.03197	0.05
90	20	20	0.8333	0.2165	0.02397	0.98
110	21	21	0.8	0.2165	0.001649	0.9
+ 110	22	–	–	–	—	—
+ 130	23		–	–	—	–
145	24	24	0.5	0.1082	0.0098	0.02
160	25	25	0			

PRACTICAL - 4

Aim :→ To compute the estimate of the survival function & standard error of estimated survival function.

Experiment :→

Suppose that 22 paitients survive time with solid tumour was seam as following :→

S. No.	Survival time	S No.	Survival time
1	3.5	3	18
12	5	13	20
7	6.5	2	+8.4
8	6.5	4	+9
5	7	21	+8
6	7	14	+9.5
17	8	15	+11
22	10	18	+12.5
9	12	16	+13
10	15	19	+15
11	16	20	+17

Write the procedure to estimate the survival & S.E of the estimated survival. Also a plot the estimated probilty density function.

Theory & FORMULA :→

This method is extensively used life table meth in which ungrouped survival time are given & based on type I progressively censored sample.

$$p_i = P(T > t_i \mid T > t_{i-1})$$

$$p_i = \frac{n-r}{n-r+1} \qquad \hat{S}(t_K) = \prod_{i=1}^{K} \frac{n-r}{n-r+1}$$

$$Var(\hat{S}(t_K)) = (\hat{S}(t_K))^2 \cdot \frac{1}{(n-r)(n-r+1)}$$

Theory & FORMULA :→

This method is an extension of life table method in which ungrouped survival times are given & is based on type I progressively censored sample.

$$p_i = P(T > t_i \mid T > t_{i-1})$$

$$p_i = \frac{n-r}{n-r+1} \qquad S(\hat{t}_k) = \prod_{i=1}^{k} \frac{n-r}{n-r+1}$$

$$Var(S(\hat{t}_k)) = S(t_k) = \frac{1}{(n-r)(n+1)}.$$

CALCULATION :→

Arranged the data in ascending order with respect to flow.

S.No.	Survival time	P_i	r_i	p_i	$S(T_k)$	$Var(S(T_k))$	S.E
1	3.5	1	1	0.9545	0.9545	0.019	0.444
12	5	2	2	0.9923	0.9091	0.0037	0.612
7	6.5	3	3	0.95	0.8636	0.0053	0.072
8	6.5	4	4	0.94774	0.8182.	0.67	0.08
5	7	5	5	0.9445	0.7227	0.79	0.69
7	7	6	6	0.9412	0.7272	0.0783	0.09
17	8	7	7	0.9375	0.6818	0.0098	0.09
21	+8	8	—	—	—	—	—
2	+8.4	9	—	—	—	—	—
4	+9	10	—	—	—	—	—
14	+9.5	11	—	—	—	—	—
22	10	12	12	0.9091	0.8198	0.0114	0.0
15	+11	13	—	—	—	—	—
9	+12	14	14	0.8889	0.5509	0.0134	0.1
18	+12.5	15	—	—	—	—	—
16	+13	16	—	—	—	—	—
10	15	17	17	0.8333	0.4591	0.01634	0.
19	+15	18	—	—	—	—	—
11	16	19	19	0.75	0.3443	6.01907	0.1
20	16	20	—	—	—	—	—
3	+17	21	21	0.5	0.1772	0.01991	0.1
13	280	22	22				

PRACTICAL - 05

Aim :→ To compute hazard fn.

Problem :→ A study was conducted diabetic patient on private clinic & their survival time (in year) follow weibull distribution & plot hazard function following parameter

t (years)	PDF	CDF
0.95	0.07	0.003
5.18	0.077	0.176
5.74	0.085	0.221
6.31	0.091	0.27
6.52	0.094	0.2195
7.78	0.101	0.414
8.97	0.099	0.558
9.92	0.098	0.56
9.24	0.098	0.586
9.57	0.096	0.742
11.3	0.77	0.781
11.66	0.073	0.789
11.9	0.073	0.787
12.65	0.069	0.833
13.34	0.049	0.872

Theory & formula :→

$S(t) = P(T>t) = P$ (an individual survives longer than t).

$= P(T>t) = 1 - P(T<t) = 1 - F(t)$

$h(t) = f(t) / S(t)$.

Weibull Distribution
Characterised by 2 parameters γ & λ
γ → determines shape, λ → determines scale
$h(t)$ → remains constant when $\gamma = 1$, increases when $\gamma > 1$ & decreases when $\gamma < 1$ as t increases.

$$f(t) = \lambda\gamma\,(\lambda t)^{\gamma-1}\, e^{-(\lambda t)^{\gamma}}$$

$$F(t) = 1 - e^{-(\lambda t)^{\gamma}}.$$

$$S(t) = e^{-(\lambda t)^{\gamma}}$$

$$h(t) = \lambda\gamma(\lambda t)^{\gamma-1}.$$

CALCULATION :→.

t	$f(t)$	$F(t)$	$S(t)$	$h(t)$
0.95	0.07	0.003	0.997	0.07021
5.18	0.077	0.176	0.824	0.093442
5.74	0.085	0.221	0.779	0.109114
6.31	0.091	0.271	0.729	0.124829
6.57	0.094	0.2595	0.7405	0.128941
7.78	0.101	0.414	0.586	0.172355
8.97	0.099	0.533	0.467	0.21191
9.22	0.098	0.558	0.442	0.221719
9.24	0.098	0.56	0.44	0.22277
9.54	0.096	0.586	0.444	0.231884
11.3	0.071	0.743	0.257	0.299611
11.66	0.073	0.769	0.231	0.316017
11.9	0.069	0.787	0.231	0.32394
12.65	0.059	0.833	0.167	0.35329
13.34	0.049	0.878	0.128	0.38281

RESULT :→

We calculated hazard function in calculations & plotted it on graph. Initially hazard function is increasing in slow rate then it with constant ~~rap~~ rate.

PRACTICAL - 06

Aim:→ To compute crude probability of death & partial ~~partiality~~ probability of death.

Experiment :→ Given that 1200 patient are suffering from risk diabities (r_1) lung cancer (r_2) & accidents (r_3) compute crude probability of death due r_1 & probability of related to r_3 After eliminating r_2

Age Interval	No. of patient alive at begining of interval	No. of dying in interval (d_i)
25-35	1200	105
35-45	1095	70
45-55	1025	55
55-65	970	55

No. of death due to r_1	No. of death due to r_2	No. of death due to r_3
20	20	16
8	26	14
25	15	5
27	18	18.

Theory & Formula :→

Life table :→ Let suppose that a study star with a sample of n patients and the study p is $(0, t)$ is divided into finite no. Sat k, of sub interval of equal length, generally ta to 1 year

Let. $I_i = (t_i, t_{i+1})$; $i = 1, 2 - - - k$.
be the i^{th} sub interval

Survival time corresponding to n_i will be censored survival time.

$P_i = P[T > t_i \mid T > d_{i-1}]$

$S(t_k) = P(T > t_k) = \prod_{i=1}^{k} P_i$

Here,

$u_i, w_i = 0$

Then, $\hat{q}_i$ = P[dying in I_i | survival at the begining of the interval]

$\hat{q}_i = \frac{d_i}{n_i}$, $\hat{P}_i = 1 - \frac{d_i}{n_i}$

$Q_{i\delta}$ → crude prob of death

↳ It is the prob of an individual is alive at the time u_i will die in the interval u_i, u_{i+1} due to cause δ, in the presence of all the risk.

operating in the population

$\sum_{i=1}^{k} Q_{i\delta} = q_i. \sum_{\delta=1}^{n} d_{i\delta} = d_i$

Estimated Survival rate.

$\hat{\lambda} = 1/\mu$

Estimated S.E of mean survival rate

$= \sqrt{Var(\hat{\mu})} = \frac{\hat{\mu}}{\sqrt{d}}$.

CALCULATION :→

$n = 20$

$t_1 = 5$
$t_2 = 6$
$t_3 = 7$
$t_4 = 8$
$t_5 = 9$
$t_6 = 10$
$t_7 = 12$
$t_8 = 15$
$t_9 = 20$
$t_{10} = 21$
$t_{11} = 25 = t(d)$

$\sum t_i = 138$

$\hat{u} = \frac{1}{\lambda} = 33$

$\hat{\lambda} = 0.0303$

$SE = \sqrt{Var(u)} = 9.9498.$

Result :→

Estimated mean survivaltime $\hat{u} = 33$.

Estimate mean survival rate $\hat{\lambda} = 0.0303$

Estimated SE of MSR $= \sqrt{Var(u)} = 9.9498$

PRACTICAL - 7

Aim:→ To explain type of Cencoring used in the given data and also to estimate mean survival ~~rate~~ time, SE of mean survival rate time.

Experiment :→

Suppose that in a study of efficiency of new drug 20 paitient with tumor are given the drug the experiment decided to terminate the study after the death of 11 paitient. The survival time in a week. 5, 6, 7, 8, 9, 10, 12, 15, 20, 21, 25. Assuming the time of death 3 paitient follow exponential distribution. Explain applied cencoring scheme in detail & estimate mean survival rate & SE of mean survival rate.

Theory & Formula

Type II Cencoring

Suppose that study starts with a sample of n patient and is terminated as soon a fixed no. of survival time is recorded

i.e. $t_{(1)} < t_{(2)} < \ldots < t_d$

Estimated mean survival time $t \sim \exp$

$$\hat{\mu} = \frac{1}{\hat{\lambda}} = \frac{\sum_{i=1}^{d} t_i + (n-d) t_d}{d}$$

$$\hat{\theta}_{ix} = \hat{P}_1 \frac{d_{ix}}{n_{i+1}} \quad \& \quad \hat{P}_1 = \frac{n_{i+1}}{n_i}$$

where $n_{i+1} = n_i - d_i$

$$\hat{\theta}_{is} = \frac{d_{is}}{n_i}$$

Partial crude probability of death.

$$Q_{i\cdot 8 6} = \frac{d_{is}}{d_i - d_{ic}}\left(1 - \left(\frac{l_{i+1}}{l_i}\right)^{\left(1 - \frac{d_{ic}}{d_i}\right)}\right)$$

where $8 = R_3$, $6 = R$.

Result

Partial probability of death due to r_1 & probability of death due to r_3 after eliminating r_1 is in calculation table

PRACTICAL 8

Aim:→ To compute estimated survival time & variance of estimated survival time.

Experiment :→

For the following data compute estimated survival time & variance estimated survival rate time using appropriate method & interprate the result.

Time four-diagnosis	No. of person lost to follow up	No. of withdrawal	No. of dying	No. of entity
0-5	18	0	731	949
5-10	16	0	52	200
10-15	8	67	14	132
15-20	0	33	10	43

Theory & Formula :→

Let I_i be the ith subinterval where {i = 1, 2, 3 -

$I_i = (t_{i-1}, t_i)$

n_i = no. of person surviving at the begining of I_i

d_i = no. of death in the interval (t_{i-1}, t_i)

l_i = no. of person lost to follow up.

w_i = no. of person widthdraw theirng

n_i' = No. of person exposed to a risk of death.

Survival time corresponding to n_i will be cencored survival time.

$$S(d_k) = \prod_{i=1}^{k} p_i$$

If $l_i \neq 0$, $w_i \neq 0$, then,

$$n_i' = n_i - 1/2(l_i + w_i)$$

$$\hat{q}_i = \frac{d_i}{n_i'} , \quad \hat{p}_i = 1 - \frac{d_i}{n_i'}$$

CALCULATION:→

Time for diagonis	No. of person lost to follow up (l_i)	No. of widrawal	No. of dying d_i	No. of entry n.	n_i'	$\hat{q}_i$
0-5	18	0	31	949	940	0.7766
5-10	16	0	52	200	192	0.0083
10-15	8	67	14	132	94.5	0.14899
15-20	0	33	10	43	26.5	0.377358

P_i	$S(d_K)$	$Var(S(\hat{d}_K))$
0.22234	0.22234	0.00183942
0.729167	0.162133	0.000148646
0.851852	0.138105	0.000142966
0.622642	0.08599	0.000224534.

$$\hat{S(t_k)} = \prod_{i=1}^{k} \left(1 - \frac{d_i}{n_i}\right)$$

$$var(S(\hat{t_k})) = [S(t)]^2 \sum_{i=1}^{k} \frac{d_i}{n_i(n_i - d_i)}$$

RESULT: →

PRACTICAL - 09

Aim :- To compute the crude probability of death

Experiment :- For the pandemic situation the following data given us the death due to risks (C, 1D, -19) Road accident, CVD, Thalasemia

Age interval	Mid Year population	Death due to Covid19	Death due to RA	Death CVD	Death thalo	
1-5	70630	740	60	615	28	5706
5-10	81911	440	80	819	15	6700
10-20	74885	384	250	1084	77	7000
20-35	105675	3050	300	1250	58	15000
35-50	127567	3500	70	1157	60	16000
50-70	140752	3623	58	1652	61	15779

Find the crude probability of death due to Covid 19 & Thalassemia

Theory :- Crude probability of death is the probability of dying due to specific cause when all the other risk are present in the population

It is denoted by Q_{is}

$$Q_{is} = \frac{d_{is}}{l_i}$$

where, d_{is} → No. of deaths due to risks R_s
l_i → Number of survivals.

Calculations :→

Age	d_{i1}	d_{i2}	l_i	$Q_{i1} = d_{i1}$	Q_{i2}
1-5	740	28	70630	0.00537	0.0080
5-10	440	15	81911	0.01048	0.001
10-20	384	77	74885	0.02886	0.001
20-35	3050	58	105675	0.02744	0.001
35-50	3500	60	127567	0.00573	0.00
50-70	3623	61	140752	0.02574	0.001

Results :→

The crude probability of death due to two risk are given by

Age	Covid-19	Thalassemia
1-5	0.01048	0.0003
5-10	0.00537	0.001
10-20	0.00513	0.0010
20-35	0.02886	0.005
35-50	0.02744	0.004
50-70	0.02574	0.0043

PRACTICAL - 10

Aim :→ To compute Type I & Type II censoring.

Experiment :→ Two experiment started their study with 23 paitients suffering from tumor. First experiment decided to terminate the study after 1 month 130 days, & second terminates his study after death of 9 paitients.

Patient No.	Survival Time (days)	Patient No.	Survival function
9	7.5	14	40
11	8	18	36
1	10.5	15	30
7	13	23	25
10	15.5	22	34
12	16	8	28
16	18	4	37
17	25	20	29
19	26	21	40
13	30	03	43
6	32	05	55
2	37		

Assuming that the time of death of these patients to exponential distributed. Explain censoring scheme used in both experiment. Compute estimated mean. Survival time.

Theory :→

In type I progressively censored data the no. of patients to be censored or observed in a random variable. But the time of study is fixed although it varies from patients to patient

$$f(t) = \lambda e^{-\lambda t} \quad , \lambda \geq 0, t \geq 0$$

$$E(T) = \mu = 1/\lambda$$

$$\hat{\lambda} = \frac{\Sigma \delta i}{\Sigma ti \delta i + \Sigma Ti(1-\delta i)} \quad \hat{\mu} =$$

$\Sigma \delta i$ → no. of uncensored observation

Type II censored data

Suppose that a sample of n patient enter the study at same time & study is terminated as soon as (d) d-th ($d < n$) occurred

$$f(t) = \lambda e^{-\lambda t}$$

$$\hat{\lambda} = \frac{d}{\sum_{i=1}^{n} \delta_i t_i - (n-d) t(d)}.$$

$$\hat{\mu} = \frac{1}{\hat{\lambda}}$$

$t_i \rightarrow$ Survival time of i^{th} patient

$\delta_i \rightarrow$ Indicator variable

$\delta_i \rightarrow \begin{cases} 1, & \text{If observation is uncensored} \\ 0, & \text{otherwise.} \quad t_i < T_c \end{cases}$

CALCULATON :→

(1) Type I Censoring

Patient Number	Survival Time (days)	Patient No.	Survival Time (in days)
9	7.5	14	30+
11	8	18	30+
1	10.5	15	30+
7	13	23	25
10	15.5	22	30+
12	16	8	28
16	18	4	30+
17	25	20	29
19	26	21	30+
13	30	3	30+
6	30+	5	30+
2	30+		

Study is terminated After 1 month = 30 days

Patient No.	Survival time	δ_i
9	7.5	1
11	8	1
1	10.5	1
7	13	1
10	15.5	1
12	16	1
18	18	1

17	25	1
23	25	1
19	21	1
8	28	1
20	29	1
13	30	1
15	30	1

$$\hat{\lambda} = \frac{\sum \delta_i}{\sum t_i \delta_i + \sum T_i (1-\delta_i)}$$

$\sum \delta_i = d = 14$

$$= \frac{14}{281.5 + (23-14) \times 30} = \frac{14}{551.5} = 0.254.$$

$$\hat{\mu} = \frac{1}{\hat{\lambda}} = 39.39.$$

(ii) Study is terminated After the death of 9 paitient

Patient No.	Survival time	δ_i
9	7.5	1
11	8	1
1	10.5	1
7	13	1
10	15.5	1
12	16	1
16	18	1
18	25	1
23	25	1

$$\hat{\lambda} = \frac{nd}{\sum_{i=1} \delta_i t_i - (n-d)\, t(d)}$$

$d = 9$

$\sum t_i \delta_i = 138.5$

$(n-d) = 23-9 = 14.$

$t(d) = 25$

$$\hat{\lambda} = \frac{9}{13.85 + 14 \times 25}$$

$$\hat{\lambda} = 0.018423.$$

$$\frac{1}{\mu} = \hat{\lambda} = 54.277$$

Result :→

(i) The mean survival time of Type I Censoring is 39.39 days.

(ii) The mean survival time of Type II. Censoring is 54.277 days.

PRACTICAL - 11

Aim :→ To determine the gene frequency P_A, P_B, P_O

Problem :→ Consider the following data on the A,B,O blood group with observed frequency

Genotype :	AA	AB	BB	BO	AB	OO
Frequency :	0.08	0.30	0.04	0.12	0.06	0.40

determine the gene frequency.

Theory of Formula :→

Genotype — A_1A_1 Homozygote, $A_1A_2(A_2A_1)$ Heterozygote, A_2A_2 Homozygote

$P_{11} \rightarrow$ frequency of A_1A_1

$P_{12} \rightarrow$ frequency of A_1A_2

$P_{22} \rightarrow$ frequency of A_2A_2.

$P_{11} = \frac{N_{11}}{N}$ $N_{11} \rightarrow$ No. of individual having A_{11}

$N_{22} \rightarrow$ " " " " " A_{22}

$N_{12} \rightarrow$ No. " " " " A_{12}

$2P_{12} = \frac{N_{12}}{N}$ $P_{22} = \frac{N_{22}}{N}$

$P_{12} + 2P_{12} + P_{22} = 1$ | $A_1 = P_{11} + P_{12}$, $A_2 = P_{22} + P_{12}$

P_1 = Genotype of $A_1 = \frac{N_{11}}{N} + \frac{1}{2}\frac{N_{12}}{N}$

P_2 = Genotype of $A_2 = \frac{N_{22}}{N} + \frac{1}{2}\frac{N_{12}}{N}$

With $P_1 + P_2 = 1$

CALCULATIONS:→ Locus blood Group.

A (P_A) B (P_B) O (P_O)

$$P_A = P_{AA} + P_{AO} + P_{AB}$$

$$= 0.08 + \frac{0.30}{2} + \frac{0.06}{2}.$$

$$= 0.08 + 0.15 + 0.03 = 0.26$$

$$P_B = P_{BB} + P_{AB} + P_{BO}$$

$$= 0.04 + \frac{0.12}{2} + \frac{0.06}{2}$$

$$= 0.04 + 0.06 + 0.03 = 0.13.$$

$$P_O = P_{OO} + P_{BO} + P_{AO}.$$

$$= 0.40 + \frac{0.12}{2} + \frac{0.3}{2}.$$

$$= 0.4 + \frac{0.12}{2} + \frac{0.3}{2}.$$

$$= 0.4 + 0.06 + 0.05 -$$

$$= 0.61$$

RESULT:→

P_A Gene frequency → 0.26

P_B Gene frequency → 0.13

P_O Gene frequency → 0.61

PRACTICAL-12

Aim:→ To compute probability in genetics.

Experiment → for the gamets AB, Ab, aB, ab with

$g_1 = P(AB)$
$g_2 = P(Ab)$
$g_3 = P(aB)$
$g_4 = P(ab)$

We have the following values

$g_1 = 0.3$
$g_2 = 0.4$
$g_3 = 0.1$
$g_4 = 0.2$

$\lambda = 0.4$.

Compute g_1, g_2, g_3, g_4 for $1^{st}, 3^{rd}, 5^{th}$ & 7^{th} generation respectively.

Theory :→

Probability of gamete in nth generation is given by.

$g_1^{(n)} = g_1 - [1-(1-\lambda)^n]\Delta_0$
$g_2^{(n)} = g_2 + [1-(1-\lambda)^n]\Delta_0$
$g_3^{(n)} = g_3 + [1-(1-\lambda)^n]\Delta_0$
$g_4^{(n)} = g_4 - [1-(1-\lambda)^n]\Delta_0$

where $\Delta_0 = \begin{vmatrix} g_1 & g_2 \\ g_3 & g_4 \end{vmatrix}$ such that $g_1 + g_2 + g_3 + g_4 = 1$

Calculation :→

$$\Delta_0 = \begin{vmatrix} 0.3 & 0.4 \\ 0.1 & 0.2 \end{vmatrix}$$

$$= 0.06 - 0.04 = 0.02$$

$g_1^{(1)} = 0.3 - [1-(1-0.4)]0.02$
$= 0.3 - [1 - 0.02] = 0.292$

$g_2^{(3)} = 0.4 + [1-(1-0.4)^3] \times 0.02$
$= 0.4157$

$g_3^{(5)} = 0.1 + [1-(1-0.4)^5] \times 0.02.$
$= 0.1184.$

$g_4^{(7)} = 0.2 - [1-(1-0.4)^7] \times 0.02.$
$= 0.1806$

RESULTS :→

We have,

Probability of gamete AB in 1st generation = 0.292
Probability of gamete AB in 3rd generation = 0.4157
Probability of gamete in AB 5th generation = 0.118
Probability of gamete in 7th generation = 0.186

F. Ay
15/5/24