	Date :
	VIT2502 - Data Analytics & Visrialization
	Assignment - 2
	J
	BY: R. NITHYASRI
	REG NO: 3122 22 5002 086
	CLASS : IT - 'B'.
	What is the difference between censored and uncersored data in survival analysis? Provide examples of each.
	uncersored data in survival analysis?
	Provide examples of each.
1)	Censored data: Censored data occurs when we
	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
	for some subjects but we do know that
	the survived on persisted up to a certain
-	for some subjects, but we do know that they survived on persisted up to a certain point This resulty happens for one of the
4	following gleasone:
-	20
	Right Censoring: - It happens when the subject hasn't experienced the event by the end of the study pariod on has dropped out before the study ends.
1	hasn't experienced the event by the end
	of the study period on has dropped out
	before the study ends.
3 3 3 4 5	heft censoring: When we know the event happened before a certain time but don't know exactly when.
1	happened before a certain time but don't
1	know exactly when.
1	O .
	Interval consoring: When the event happens
5	within a known time interval but
5	Interval censoring: When the event happens within a known time interval but not at an exact time.
9	
2	
374 24 07 3	

	Date :	
	To all the second of the secon	
	The A Clinical Man M. a police F dogs	
	not relapse before the study ends,	
	their time is reconded up to the	79
	study's end date, the data is right -	
	censored.	
1		
-		
	If a participant arops out of a study for	
	personal reasons, Their data would also	
	be right - consored since we only know	
	If a participant drops out of a study for personal reasons, their data would also be right - censored since we only know they survived the dropout date.	dy
1	2) Uncersored data: We know the exact time at which the event occurred for	
1	to turkich the event occurred for	N
-	Time at the record free	
1	a subject.	
-	a contra substantia de la companya della companya della companya de la companya della companya d	
	Eg: - In a study tracking time to recovery	-
	after a surgery if a patient fully	•
	recovers within the study period	4
1	Eg:-In a study tracking time to recovery after a swigery, if a patient fully recovers within the study period and the exact recovery time is recorded	-
1		-
-	The state of the s	7
1	In a job satisfaction study tracking	-
	employee twenover, if an employee leaves	-
	the company during the observation	-
Ì	beried their exact tenure diviation	,
ľ	is rencens ored data.	
+	US WILLEAS OF CERTIFICATION OF THE PROPERTY OF	
-	The state of the second control of the second of the secon	-
		_
1	the state of the second of the	_

,

9		
6		
ê		Date :
-	9.5	Charles
	2)	Survival Functions:
5		Explain the Kaplan-Meier estimator and its
		Explain the Kaplan-Meier estimator and its use in estimating the survival function How does it handle consorred data?
		How does it handle consored data?
		The Kaplan-Meier estimator is a non-
9-		The restrict - state of the sta
		parametric tool used in survival
		analysis to estimate the survival function
		The Probability of the second
		subject survives beyond a given time.
		It is particularly useful for tarding
		subject survives beyond a given time t. It is particularly useful for handling right - censored data.
-		o .
	, =	Key points:
	-19	The state of the s
	۵۱۵	1) Calculation: - The estimator calculates survival
		probabilities at each event time ti
		using only those subjects who are still
		Description: The estimator calculates survival probabilities at each event time ti, using only those subjects who are still 'at risk'.
		2) Censored Data Handling: - Censored subjects
		2) Censored Data Handling: - Censored subjects Contribute to the rusks set until the
		time they are consorred.
-		
		3) Kaplan - Meier Curve: - The results are
		visualized as a stepwise curve that decreases
		at each event time, with small vertical ticks marking censored data
	1.	vertical ticks marking censored data
		points.
3		
2		

,	
	Date :
	Kaplan - Meier Survival Function:
1	S(+) = TT (1-di)
	$S(+) = \prod_{i \leq +} \left(1 - \frac{di}{ni} \right)$
	Kaplan - Meier Survival Function:- S(+) = TT (1 - di) ti = ti
71 5	ti - Time of the ith event
	di = Number of events at ti
	ai = Mumber of
~ 1 - 6	ni = Number of subjects at risk before time ti.
	ni = Number of subjects at risk before time ti.
	Deforte Terrer
2)	Hard Hazard functions:
3)	
	The hazard function, h(+) represents an instantaneous rate at which events
	The harmon grate at which events
	at a prosticular time i, giver
- 12-11-14	that the individual has survived up
* ,	that The sharves !
ê - [*]	to that time
	h(+) = lim Pg(+≤T<++Δ+1T>t)
	nci) = xim : sc = 1
	$\Delta t \rightarrow 0$
<u> </u>	1 A A A A A A A A A A A A A A A A A A A
	lelationship to survival function,
	- J + (u)du
ur i	SCH) Marie Gielle de la
	als table second second second second as a second second
	A higher hazard implies higher risk of the event occurring.
لمرائد	the event occurring.

	Date :
	A constant hazard suggests a steady risk over time.
1	
1	A decreasing hazard over time might
	A decreasing hazard over time might occur in scenarios where early failures are more common but those who survive
	are more common but those who survive
	initially face reduced risk later.
(د	Propositional hazards:
	A COLUMN TO STATE OF THE STATE
	The proportional hazarde assumption in the
	Cox Proportional Hazards model states
	that the hazard ratio between individuals
	with different covariates is constant over
	time. This means that the effect of each
	predictor on the hazard remains the
5	same throughout the study period.
_ =	Ti de esumption:
	Testing for the Assumption:
	1) Schoenfeld Residuals
	2) Log-Log Survival Curves
	3) Time - Dependent Covariates.

	. }		16
			V
		Date :	
1	F7		V
	2)	Time-varying covariates:	
		Time in a majurius and usu	
		Time-varying covariates in survival analysis	-
		Can be allowing covariates to change	
. n	* 4	over time, typically through an extended	9
- 1	1 1 1	Cox model or sphilling include increased	See a
		model complexity difficulty in interpreting	
		Cox model or splitting time into intervals. Challenges include increased model complexity, difficulty in interpreting hazard ratios, and potential violations.	-
		of the propositional hazards assumptions.	
	,		~
	2 2 3	Relation between hazard function and	
		Covariates:	
	. 0	β, X; (+) + β, X; (+) ····	
		hi(+) = ho(+) e	
	- 1 ·	Liener braker ett op value of	6
		where Xij (+) represents the value of	5
		covariate jat time + for individual	65
			5
		Lingling of the forest of the	3
		promote landinale pul soletier	
		Established the bounded in Line 18	-
			-
			-4
			-
			1
			1