

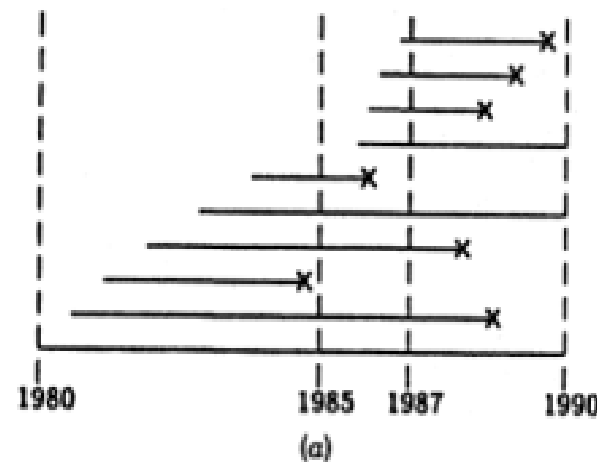
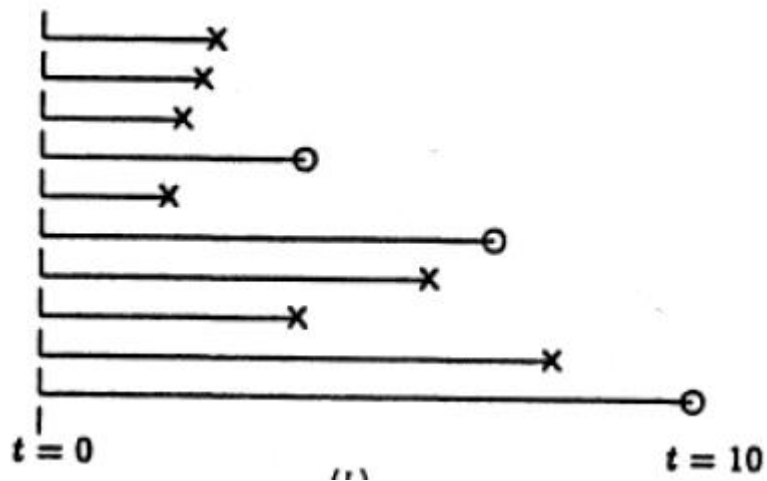
Túlélés analízis

(SURVIVAL ANALYSIS)

Története:

- ***Aktuárius***: számszerűsített válaszokat ad a kockázatok pénzügyi hatásainak kezelésére: a) Aktuáriások (Berkson és Gage 1950), b) Cutler és Ederer fejlesztette tovább az eljárást (1958).
- Egyéb alkalmazások pl. orvostudomány.
- A módszer célja a túlélési függvény becslése, ill. az ehhez kapcsolódó szignifikanciavizsgálatok.
- A túlélési függvény megadja, hogy egy kezdő eseménytől számítva várhatóan hogyan csökken a túlélők százalékos aránya az idő függvényében.

- **Vizsgált célváltozó:** egy adott kiindulási időponttól (diagnózis, bevásárlás, valamilyen beavatkozás) egy meghatározott eseményig (legtöbbször halálozás, de lehet más jól meghatározott esemény időpontja) eltelt idő, más néven „**túlélési idő**” (*survival time*).
- **Fő vonás:** az esetek egy részét a vizsgálók különböző okok miatt nem követik(hetik) az esemény bekövetkezéséig (*cenzorálás*).



Probléma:

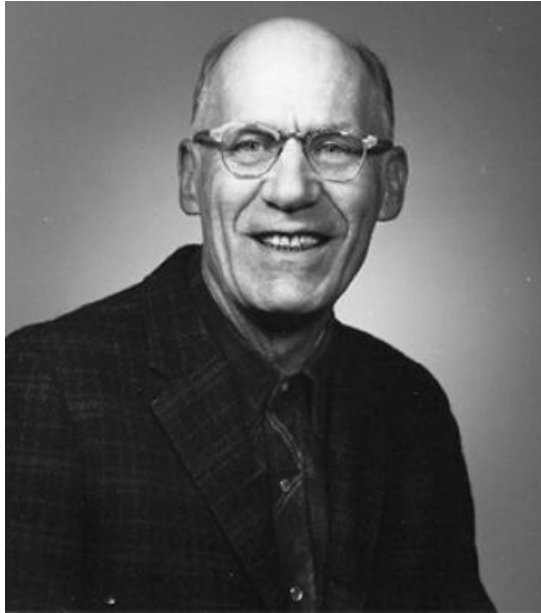
- Túlélési idő vizsgálata speciális vizsgálati módszereket igényel (pl. két csoport között az idők átlagait nem lehet direkt módon összehasonlítani)
- A túlélési idő nem normális eloszlású (korlátozó feltétel).

Definíciók:

- *Esemény (endpoint, failure)*: bármilyen megfigyelés eredménye (műtét utáni felépülés).
- *Megfigyelési idő (observation time)*: Az esetek többségében egy előre definiált időintervallum.
- *Követés (follow-up)*: A betegek sorsának figyelése.
- *Kiesett egyének (drop-out)*: a vizsgálatból kiesett egyének.
- *Visszavonás (withdrawing)*: pl. együttműködés hiánya miatt egyéneket kizárunk a vizsgálatból.
- *Túlélési idő (survival time)*: a megfigyelés kezdetétől a vizsgált eseményig eltelt idő (pl. hány nap telik el az operációt követően a felépülésig).
- *Nem teljes adat (censored data)*: olyan egyének, akik 'elvesznek' a megfigyelés során (pl. elköltöznek), de addigi adataikat felhasználjuk.

Legismertebb módon készíthetünk túlélési függvényt:

- Kaplan–Meier eljárás (product–limit módszer)
- life–table (halandósági tábla) analízis (az előző módszert is tartalmazza).



Edward Lynn Kaplan

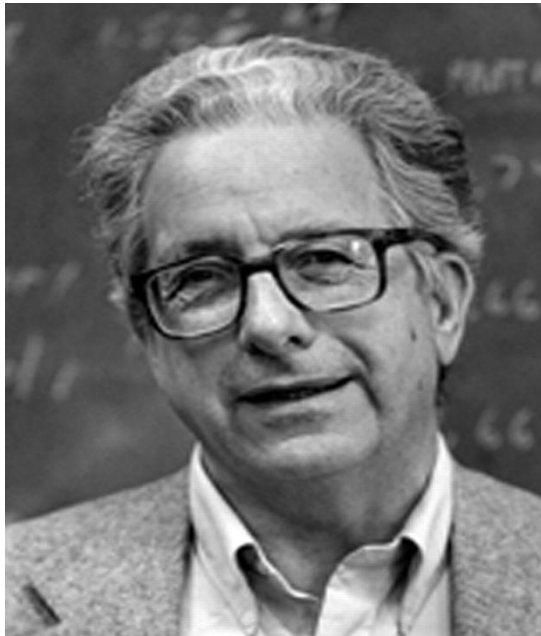
Born

May 11, 1920, Philadelphia

Died

September 26, 2006, Corvallis, Oregon

Kaplan E. L., Paul Meier. Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association*, Vol. 53, No. 282 (Jun., 1958), pp. 457-481).



Paul Meier (July 24, 1924 – August 7, 2011)

- A túlélés vizsgálatának egyik lehetséges módja a **halandósági tábla (life table)** készítése. Ez a következőképpen történik:
- A 0-tól a maximális követési időig terjedő időtartamot szakaszokra osztjuk, és meghatározzuk az egyes szakaszokra jellemző halandóság értéket, amely az abban az időszakban meghalt személyek és az időszak ***közepén*** életben levők hányadosa.
- Ezt az értéket 1-ből kivonva kapjuk az időszakra jellemző túlélési arányt. Annak valószínűsége, hogy egy adott személy valamely időszak végén életben van, az addigi időszakokra számított túlélési arányok szorzata.

Kaplan-Meier módszer

- A Kaplan-Meier (KM) módszer nagyon hasonlít a halandósági tábla módszeréhez, azzal a különbséggel, hogy a követési idő nincs szakaszokra osztva, ehelyett a kockázatot és a túlélési valószínűséget minden olyan **időpontban** meghatározzuk, amelyben legalább egy „halálozás” történt. (**pontbecslés**)
 - Kaplan E. L., Paul Meier. Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association, Vol. 53, No. 282 (Jun., 1958), pp. 457-481).

Túlélés valószínűségének becslése: Kaplan-Meier módszer

k számú esemény:

$$t_1 < t_2 < t_3 < t_4 < t_5 < \dots < t_k$$

$$S(t_j) = S(t_{j-1}) \left(1 - \frac{d_j}{n_j}\right), \text{ ahol}$$

n_j = az eseményt nem elszenvedő minták száma a t_j időpont előtt

d_j = az események száma a t_j időpontban

$t_0 = 0$ és $S(0) = 1$

Kaplan-Meier túlélési görbe: $S(t_j)$ alakulása az időben

Túlélés és hazard

Túlélés valószínűsége, túlélési függvény (Survival probability, survival function):

$S(t)$: Annak valószínűsége, hogy egy adott személy a kezdőidőponttól kezdve a t időpontig túlél.

Hazard (hazard function):

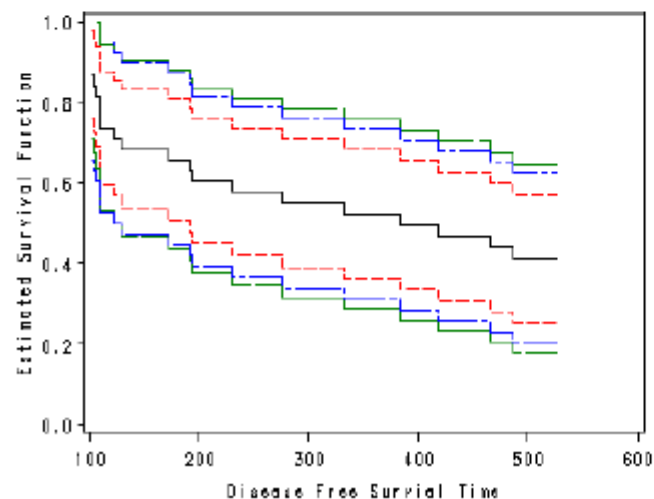
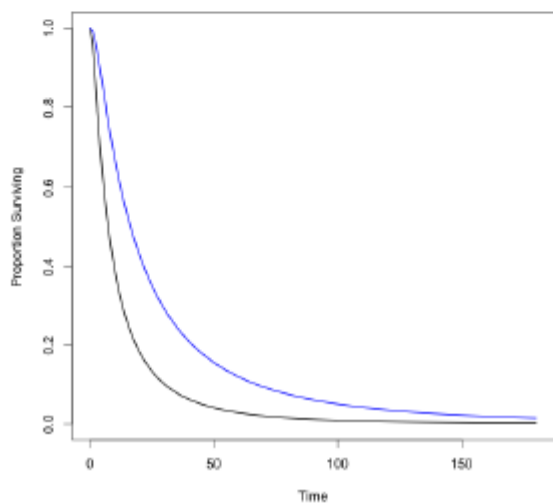
$h(t)$: Annak valószínűsége, hogy egy személy a t időpontban elszenved egy eseményt, ha addig nem szenvedett el.

Túlélési függvény

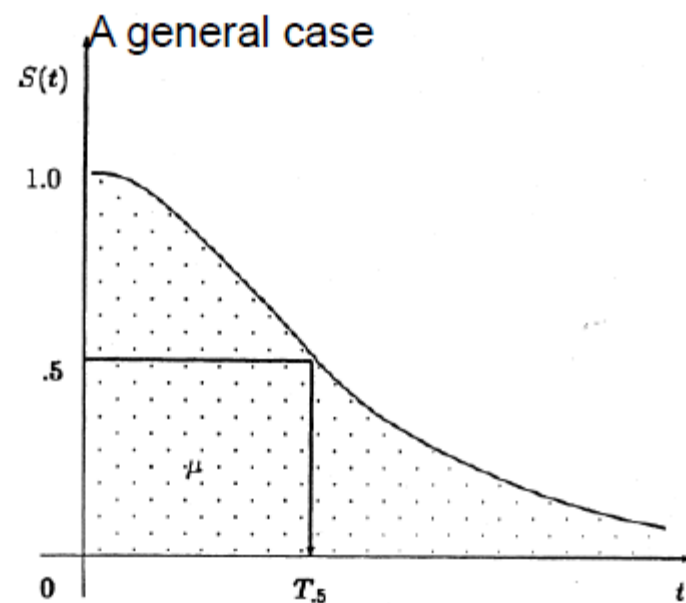
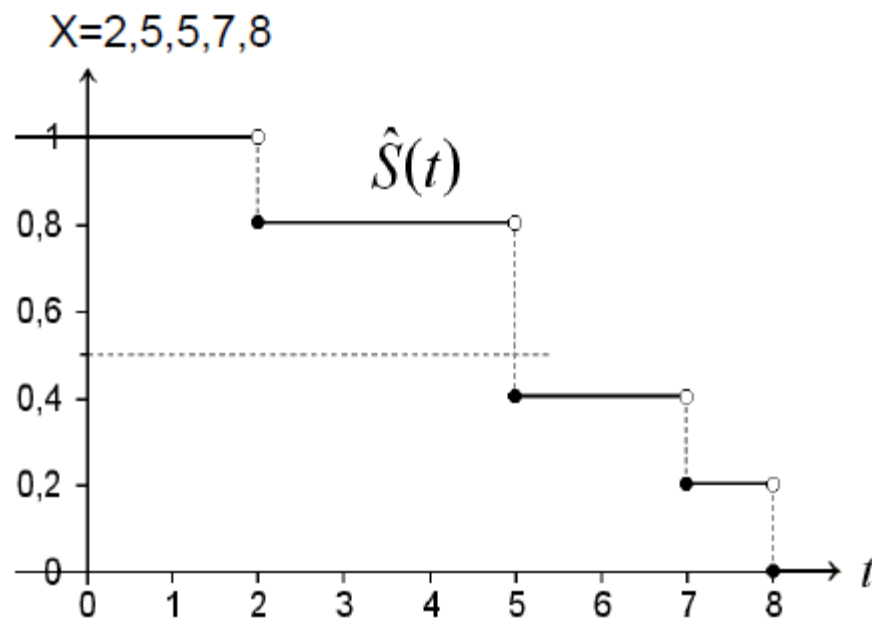
Annak valószínűsége, hogy a beteg egy adott időnél tovább él. $S(t)=P(X>t)$

Tulajdonságok:

- nem-növekvő
- $S(0)=1$ (0 időpontban (kezdetben) 1 a túlélés valószínűsége)
- $S(\infty)=0$ (ahogy időben haladunk a végtelen felé, a túlélés függvény a 0-t közelíti)
- elméletileg a túlélés függvény folytonos. A gyakorlatban az eseményeket diszkrét időpontokban figyeljük meg, így a túlélés függvényt egy lépcsősfüggvénnyel közelítjük



Medián és átlagos túlélési idő



Chap T. Le: Survival Analysis

Medián túlélés= $T_{0.5}$ az az idő, amelyben az adott betegségben szenvedő betegek fele még „életben” van
Meghatározása: az a hely a t-tengelyen, ahol a túlélési függvény „átlépi” a 0.5-es valószínűséget

Átlagos túlélés= μ (túlélési függvény alatti terület)

Egyéb vizsgálati módszerek

A ***Nelson-Aalen becslés***: egy nem paraméteres eljárás, amely a kumulatív hazard rate függvényt adja meg, cenzorát túlélési adatokból.

A Kaplan-Meier becslés általánosítható nemhomogén Markov-folyamatokra. Egy ilyen általánosítást Aalen (1978) vett figyelembe a the competing risks model-re. Különösen Aalen és Johansen szorzat-integrál formulája mutatja meg, hogy az általában Aalen-Johansen becslőnek nevezett becslés hogyan tekinthető a Kaplan-Meier becslés mátrix változatának.

Teszt statisztikák

- Standard hiba $se(\hat{s}_t) = s_t \sqrt{\sum_{i=1}^t \frac{q_i}{n_i - d_i}}$
- 95% Konfidencia intervallum $s_t \pm 1,96 * \hat{se}(s_t)$
- Greenwood – teszt
 - (khi-négyzet teszt 1 szabadságfokkal)

$$\left\{ \frac{s_t - s}{\hat{se}(s_t)} \right\}^2$$

Túlélési valószínűségek összehasonlítása

- H_0 : a két túlélési függvény azonos a két populációban,
 $S_1(t) = S_2(t)$
- H_a : a két túlélési függvény különböző a két populációban,
 $S_1(t) \neq S_2(t)$
- A több, különböző néven is elterjedt próba közti különbség a végesemények súlyozásában van.
- Ha a próba szignifikáns (pl. $p < 0,05$), akkor azt mondjuk, hogy a csoportok túlélésében szignifikáns különbség van, vagyis a csoportosító tényező (pl. rizikófaktor, vagy kezelés) szignifikánsan befolyásolja a túlélést.

Túlélési görbék összehasonlítása:

- a) *Általánosított (generalizált) Wilcoxon teszt:* a korai események nagyobb súlyt kapnak.
- b) *Mantel–Cox teszt:* egy exponenciális score teszt.. Lényegében egy log–rank teszt.
- c) *Breslow–teszt:* a korai megfigyeléseket súlyozza, kevésbé érzékeny a későbbi eseményekre.
- d) *Tarone–Ware teszt:* a Breslow és a Mantel–Cox teszt között helyezkedik el. Az eseményeket közepesen súlyozza.
- e) *Peto–Prentice teszt:* az általánosított Wilcoxon teszttel analóg.
- f) *Log-rank teszt:* a két csoport közötti mortalitási arány egymáskonstanszorosa.

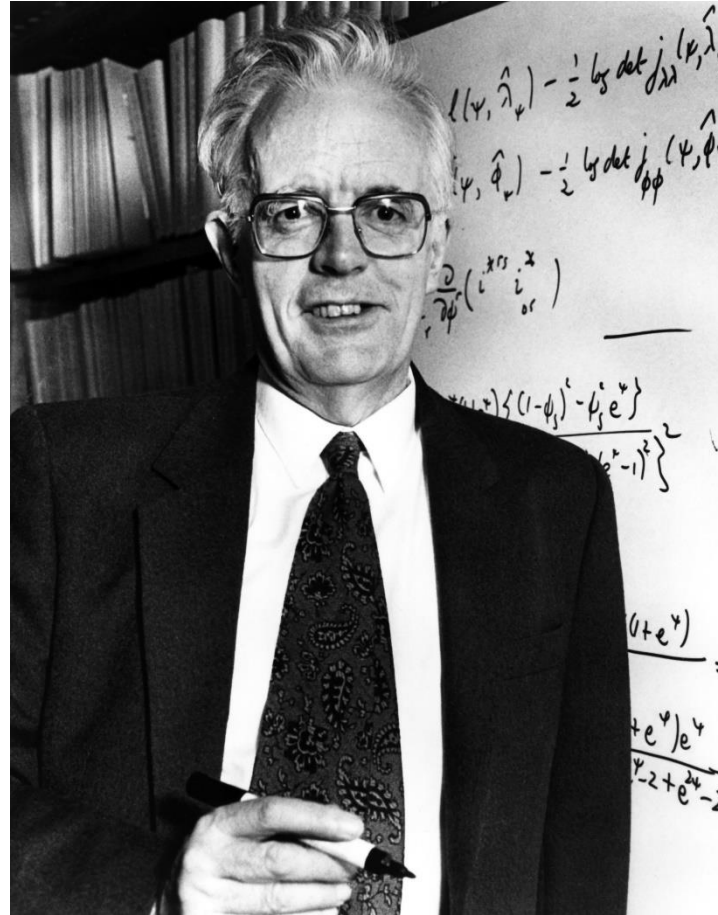
Vizsgálati meggondolások:

- 1) A vizsgálat milyen mintaszámmal, milyen hosszú ideig tartson, mi a vizsgált esemény (end point).
- 2) Hogyan kezeljük a drop out eseteket.
- 3) Ismétlődő jelenségre ne végezzünk túlélés analízist.

- 4) A túlélési kritérium, a feltétel rendszer nem változhat meg a vizsgálat folyamán.

Cox Proportional Hazards Regression Analysis

Sir David Roxbee Cox (born 15 July 1924) is a prominent British statistician.



Cox, David R (1972). "Regression Models and Life-Tables". *Journal of the Royal Statistical Society, Series B*. 34 (2): 187–220.

· Idézetek száma: 53795

Cox–regresszió

Cox (1972) a túlélési problémák vizsgálatára dolgozta ki a proportional hazards regressziós módszerét (egyszerűen Cox modell). A vizsgált esemény kockázata a hazard rate (a kumulatív túlélési görbe meredeksége egy időintervallumban). A kockázatot úgyis definiálhatjuk, mint adott t időpontban a halál bekövetkezésének valószínűségét, amikor tudjuk, hogy az egyén a t idő előtt még él. A megfogalmazást kérdés formájában is feltehetjük: a beteg milyen valószínűséggel éli meg az öt évet a beavatkozás után, ha már három évet túlél?

Ha az egyik csoportban a halál kockázata háromszorosa a másik csoporténak, akkor a kockázat állandó marad az egész vizsgálat folyamán. Ezt úgy mondjuk, hogy a két csoport hazard függvénye egymással arányos, proporcionális. Az előbbi példa kapcsán ez azt jelenti, hogy az első csoportban a halálesemény valószínűsége háromszor nagyobb, mint a másik csoportban.

A Cox–regressziós modell a vizsgált magyarázó változók relatív kockázatát becsli. A kapcsolatot mindig egy kockázatmentes (rizikófaktor mentes) csoporthoz, a baseline csoporthoz viszonyítjuk.

$$S(t) = \Pr(T > t)$$

Ez azt jelenti, hogy a túlélés függvény annak a valószínűségét mutatja, hogy a halál t koron túl következik be.

A túlélés függvény t növekedtével zéróhoz konvergál, azaz: $S(t) \rightarrow 0$ as $t \rightarrow \infty$.

A Cox–regressziós modell a vizsgált magyarázó változók relatív kockázatát becsli. A kapcsolatot mindig egy kockázatmentes (rizikófaktor mentes) csoporthoz, a baseline csoporthoz viszonyítjuk.

Jelölje $h(t)$ a kockázat mértékét a t időpontban, $\underline{x} = [x_1, x_2, x_3, \dots, x_n]$ a vizsgált magyarázó változók vektorát. A Cox–regressziós modell a kockázat nagyságát írja le a t időpontban, az \underline{x} vektor értékei mellett

$$\log h(t, \underline{x}) = \log [h_0(t)] + \sum_{i=1}^n \beta_i x_i$$

Az egyenletben a $h_0(t)$ függvény az egyén kockázatát jelenti a t időpontban, amikor a kovariánsok mindegyike 0. Az egyenlet az alábbi formában is írható

$$p(T > t, \underline{x}) = [P_0(T > t)]^{\exp\left(\sum_{i=1}^n \beta_i x_i\right)}$$

Az egyenlet a t -nél nagyobb időtartam (T) túlélési valószínűségét adja meg.

A regressziós vizsgálat legfontosabb szempontja a β_i regressziós együtthatók meghatározása. A modell lényegében logisztikus modellnek is tekinthető.

Hazárd függvény és a kumulatív hazárd függvény

A hazárd függvény, konvencionálisan λ , t időben az események aránya, feltéve ha túlélés t -ig, vagy tovább tart (azaz, $T \geq t$):

$$\lambda(t) dt = \Pr(t \leq T < t + dt | T \geq t) = \frac{f(t) dt}{S(t)} = -\frac{S'(t) dt}{S(t)},$$

A hazárd ráta/arány egy másik használatos szinonima. A hazárd függvény nem lehet negatív ($\lambda(t) \geq 0$), és az integrálja a $[0, \infty]$ tartományban végtelen kell hogy legyen, máskülönben semmi korlátozás sincs, azaz lehet csökkenő, növekvő, monoton vagy szakaszos. Egy példa a fürdőkádgörbe hazárd függvény, mely nagy, kis t értékeknél, csökken egy bizonyos minimumig, majd újra nő. Ez modellezheti néhány mechanikus rendszer tulajdonságát a meghibásodás folyamatát tekintve az egész életúton át. A hazárd függvényt a kumulatív hazárd függvény kifejezéseivel is lehet megjeleníteni (ez konvencionálisan Λ):

$$\Lambda(t) = -\log S(t)$$

$$S(t) = \exp(-\Lambda(t))$$

Cox proportional hazards regression model használatának feltételei:

1. A túlélési idők függetlensége az egyének között.
2. Multiplikatív kapcsolat a prediktorok és az esemény között.
3. Időben konstans hazard ratio megléte.

Eloszlások:

A túlélések vizsgálatokor számos eloszlással kísérleteznek, természetesen (az időskála miatt) csak azokkal, melyek a **nemnegatív valós számokon vannak értelmezve**. Ide tartozik az exponenciális, a Weibull, a Gompertz-Makeham, a lognormális, a log-logisztikus, a gamma és az inverz Gauss eloszlás is.

Az exponential regression survival model feltételezi, hogy a hazard function konstans!

Megfontolások:

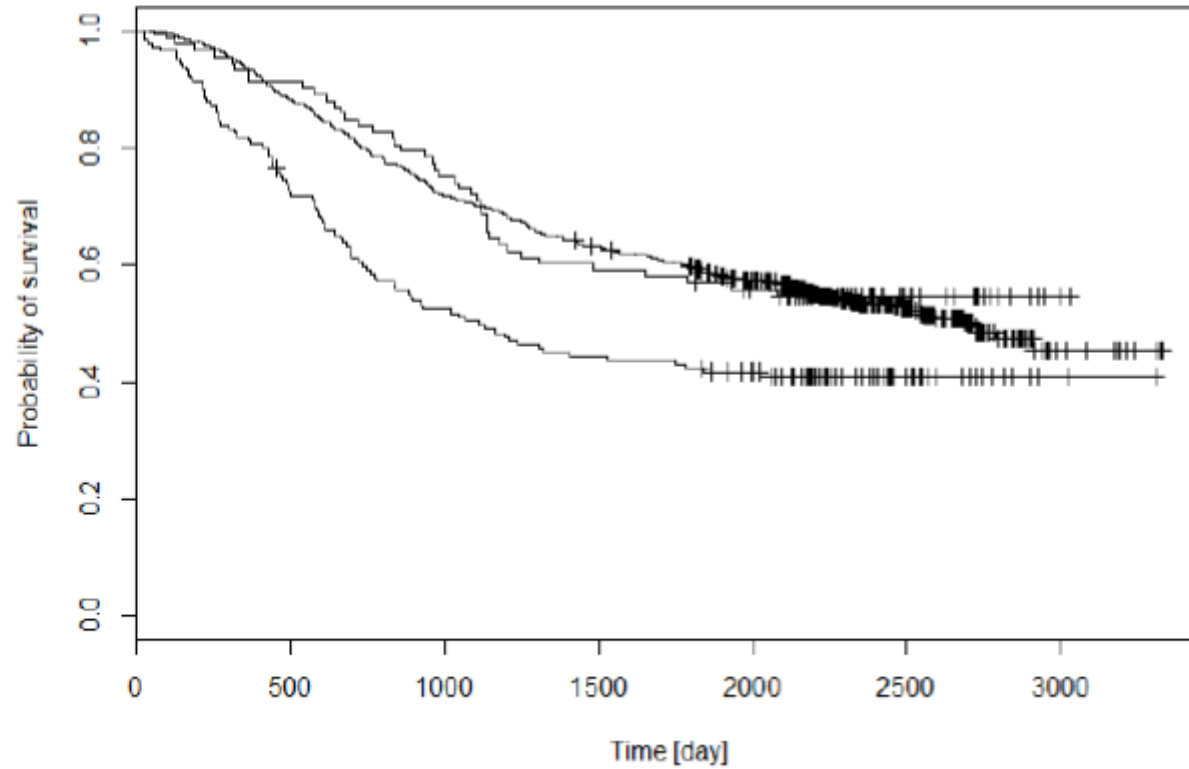
- a) Érdemes a modellvizsgálat előtt a változók függetlenségvizsgálatát elvégezni.
- b) A modell feltételezi a kockázat arányának időbeli állandóságát. (**hazard assumption**)
- c) A mintaszám megválasztásánál alkalmazzunk az ökölszabályt, hogy minden kovariánsra legalább 5 esemény jusson.
- d) A Cox-modellt gyakran alkalmazzák az exploratív vizsgálatok során hipotézisek felállítására.

Proportional hazard assumption vizsgálata

- A Cox-modell használata fontos szempont, hogy vizsgálatunk minőségbiztosított legyen, vagyis ellenőrizni kell az ún. "proportional hazard assumption" jelenséget. Ez azt jelenti, hogy ellenőrizni kell a kovariánsok hazard-jainak időbeli konstans jellegét (időfüggetlenségét), ami a modell érvényességének alapja. Erre a több lehetőség közül néhány az alábbiakban.
- A proportional hazard assumption feltétel vizsgálatakor érdemes több módszert is kipróbálni, mert akkor kapunk tisztább képet a kovariánsra nézve:
 1. Legegyszerűbb eljárás a csoportonkénti Kaplan-Meier görbék vizsgálata: ha a görbék több mint egyszer metszik egymást, akkor sérül a proportional hazard assumption.
 2. A grafikai vizsgálat kiegészíthető a Kolmogorov-féle ún. supremum teszttel is.
 3. Schoenfeld-teszt elvégzése (1982): a kovariánsok reziduálisait vizsgálja.

Arányos kockázat fennállásának tesztelése

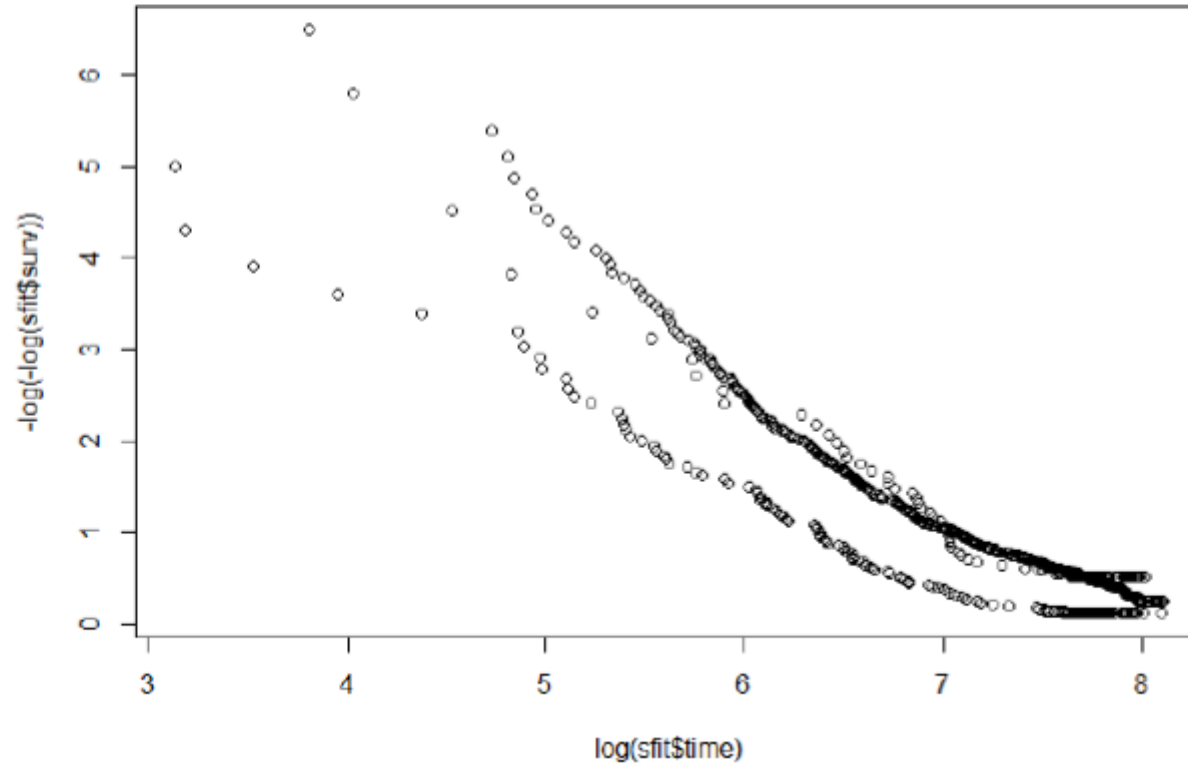
1. Túlélési görbék vizsgálata szemmel



Arányos kockázat fennállásának tesztelése

2. log-kumulatív hazard görbék vizsgálata szemmel

$\log(\text{time})$ vs. $-\log(-\log(S(t)))$



Arányos kockázat fennállásának tesztelése

3. Schoenfeld reziduálisok tesztelése

Az események bekövetkezésekor a kovariánsok megfigyelt és várható értékének különbsége.

$P < 0.05$ esetén nem teljesül az arányos kockázat feltétele!

Modell illesztés jóságának vizsgálata

Harrell's C-index (Concordance index, 1982):

- < 0.5 alatti érték nagyon rossz modellt jelent.
- 0.7 feletti értékek jó modellt jeleznek.
- $> 0,8$ feletti értékek erős modellt jeleznek.

Harrell's Concordance Statistic					
Source	Estimate	Comparable Pairs			
		Concordance	Discordance	Tied in Predictor	Tied in Time
Model	0.7134	711	283	9	2

Goodness of fit (Heart)	
Total number of cases: 65 Complete: 29 Censored: 36	
Estimation procedure converged.	
Total number	
-2 Log L	173.3592
AIC	183.3592
SBC	190.1957
R2	0.5732
Global test: All betas = 0	
Score	22.0203
Likelihood Ratio	24.6949
Wald	19.4886

Gyorsított modellek

A vizsgálatok során gyakran előfordul, hogy egy kezelés hatását szeretnénk igazolni, vagy egy új kezelést összevetni az eddig használttal. Ilyen esetekben célszerű gyorsított modelleket alkalmazni, összehasonlítva a két csoportot (kezelt-nem kezelt/új módszer-régi módszer). Fontos megjegyeznünk, hogy az egyes páciensek véletlenszerűen kerülnek a kezelt, ill. a kontrollcsoportba.

A gyorsított élet modell tehát nem más, mint a túlélésfüggvény időskála szerinti multiplikatív módosítása, vagyis a betegség sebességének változtatása. Ha $\beta = 0$, akkor a két csoport nem különbözik egymástól, $\beta > 0$ esetén károsabb az új módszer a réginél, $\beta < 0$ mellett pedig előnyös, például ha $\beta = -\log 2$, akkor a kezelés hatására az életkilátásaink romlásának sebességét felezi, vagyis hasznos.

Tobit modell

Magát a modellt (Arthur Goldberger & James Tobin, 1958) gyakran hívják cenzorált lineáris regressziós modellnek is, amelyben a magyarázó változók lineáris kapcsolatban vannak a balról (alulról) vagy a jobbról (felülről) vagy mindkét oldalról cenzorált magyarázott (vagy függő) változóval. A balról cenzorált adatok azt jelentik, hogy az adatoknak létezik egy olyan alsó korlátja, amelynél kisebb értékeket nem ismerünk. Pl. ha egy műszer alsó mérési tartománya 3 µg-ig terjed, akkor hiába van a vizsgált mintában ennél kisebb mennyiség, azt kimutatni nem tudjuk. Ilyenkor adataink balról cenzoráltak lesznek. Hasonlóan értelmezhető a felülről cenzorált adatállomány illetve a mindkét oldalról cenzorált adatállomány is. Az ilyen típusú adatok vizsgálatára hasznos a Tobit modell, szemben más esetleges megoldásokkal:

Adatelőkészítés

Data: Heart.sta (11v by 65c)

Heart transplant data from Crowley and Hu, stratified

	1	2	3	4	5	6	7	8	9	10	11
	MONTH_1	DAY_1	YEAR_1	MONTH_2	DAY_2	YEAR_2	CENSORED	AGE	ANTIGEN	MISMATCH	HOSPITAL
1	JANUARY	6	68	JANUARY	21	68	CENSORED	54	0	1.11	HILLVIEW
2	MAY	2	68	MAY	5	68	CENSORED	40	0	1.66	HILLVIEW
3	AUGUST	31	68	MAY	17	70	COMPLETE	51	0	1.32	HILLVIEW
4	AUGUST	22	68	OCTOBER	7	68	COMPLETE	42	0	0.61	ST_AND
5	SEPTEMBER	9	68	JANUARY	14	69	CENSORED	48	0	0.36	ST_AND
6	OCTOBER	5	68	DECEMBER	8	68	COMPLETE	54	0	1.89	ST_AND
7	OCTOBER	26	68	JULY	7	72	COMPLETE	54	0	0.87	BINER
8	NOVEMBER	22	68	AUGUST	29	69	COMPLETE	49	0	1.12	BINER
9	NOVEMBER	20	68	DECEMBER	13	68	CENSORED	56	0	2.05	HILLVIEW
10	FEBRUARY	15	69	FEBRUARY	25	69	COMPLETE	55	1	2.76	HILLVIEW
11	FEBRUARY	8	69	NOVEMBER	29	71	COMPLETE	43	0	1.13	BINER
12	MARCH	29	69	MAY	7	69	COMPLETE	42	0	1.38	HILLVIEW
13	APRIL	13	69	APRIL	13	71	COMPLETE	58	0	0.96	ST_AND
14	JULY	16	69	NOVEMBER	29	69	COMPLETE	52	1	1.62	ST_AND
15	MAY	22	69	APRIL	1	74	CENSORED	33	0	1.06	ST_AND
16	AUGUST	16	69	AUGUST	17	69	CENSORED	54	0	0.47	BINER
17	SEPTEMBER	3	69	DECEMBER	18	71	COMPLETE	44	0	1.58	BINER
18	SEPTEMBER	14	69	NOVEMBER	13	69	COMPLETE	64	0	0.69	HILLVIEW
19	JANUARY	16	70	APRIL	1	74	CENSORED	49	0	0.91	BINER

Data: Spreadsheet2* (7v by 65c)

Heart transplant data from Crowley and Hu, stratified

	1	2	3	4	5	6	7
	START	STOP	CENSORED	AGE	ANTIGEN	MISMATCH	HOSPITAL
1	1/6/68	1/21/68	CENSORED	54	0	1.11	HILLVIEW
2	5/2/68	5/5/68	CENSORED	40	0	1.66	HILLVIEW
3	8/31/68	5/17/70	COMPLETE	51	0	1.32	HILLVIEW
4	8/22/68	10/7/68	COMPLETE	42	0	0.61	ST_AND
5	9/9/68	1/14/69	CENSORED	48	0	0.36	ST_AND
6	10/5/68	12/8/68	COMPLETE	54	0	1.89	ST_AND
7	10/26/68	7/7/72	COMPLETE	54	0	0.87	BINER
8	11/22/68	8/29/69	COMPLETE	49	0	1.12	BINER
9	11/20/68	12/13/68	CENSORED	56	0	2.05	HILLVIEW

Start	Stop	Survival time
1/6/1968	1/21/1968	15
5/2/1968	5/5/1968	3
8/31/1968	5/17/1970	624
8/22/1968	10/7/1968	46
9/9/1968	1/14/1969	127
10/5/1968	12/8/1968	64
10/26/1968	7/7/1972	1350
11/22/1968	8/29/1969	280
11/20/1968	12/13/1968	23

Life Table készítés

Survival and Failure Time Analysis: Heart2.sta

Quick

- Life tables & Distributions
- Kaplan & Meier product-limit method
- Comparing two samples
- Comparing multiple samples
- Regression models
- Time-dependent covariates

Cancel

Options

Open Data

SELECT CASES

Life Table & Distribution of Survival Times: Heart2.sta

Raw data | Table of survival times

Variables (survival times & censoring indicator)

Survival times (1) or dates (2 or 6): none

Variable with censoring indicator: none

Code for complete responses:

Code for censored responses:

Compute table based on:

Number of intervals: 12

Stepsize (interval width): 1.

Correct intervals containing no terminations/deaths

Cancel

Options

SELECT CASES

Select survival times/dates and censoring indicator

1 - MONTH_1
2 - DAY_1
3 - YEAR_1
4 - MONTH_2
5 - DAY_2
6 - YEAR_2
7 - CENSORED
8 - AGE
9 - ANTIGEN
10 - MISMATCH
11 - HOSPITAL

1 - MONTH_1
2 - DAY_1
3 - YEAR_1
4 - MONTH_2
5 - DAY_2
6 - YEAR_2
7 - CENSORED
8 - AGE
9 - ANTIGEN
10 - MISMATCH
11 - HOSPITAL

Survival times (1), dates (2 or 6): 1-6

Censoring indicator: 7

Show appropriate variables only

OK

Cancel

[Bundles]...

Use the "Show appropriate variables only" option to pre-screen variable lists and show categorical and continuous variables. Press F1 for more information.

Life Table & Survival Time Distribution Results: Heart.sta

Variable: number of days computed from data

Variable with censoring indicator: CENSORED

Total number of valid observations: 65

uncensored: 29 (44.62%) censored: 36 (55.38%)

Quick | Advanced | Function plots

Summary: Life table

Results for model: Exponential

Parameter estimates

Summary

Cancel

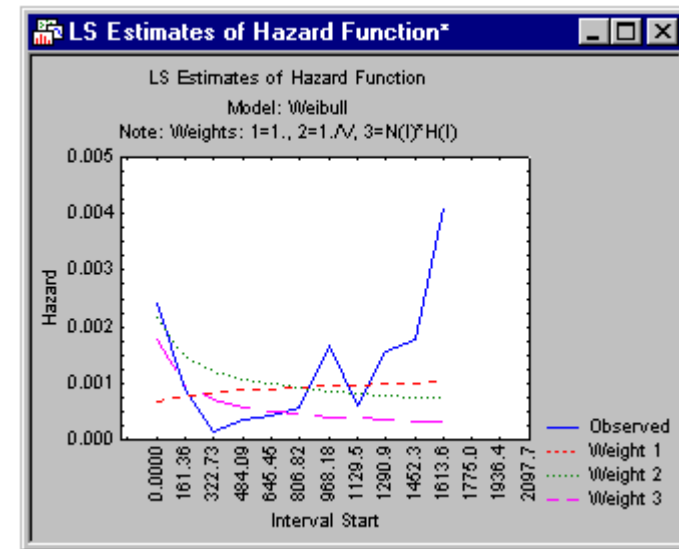
Options

Data: Life Table (Heart)*

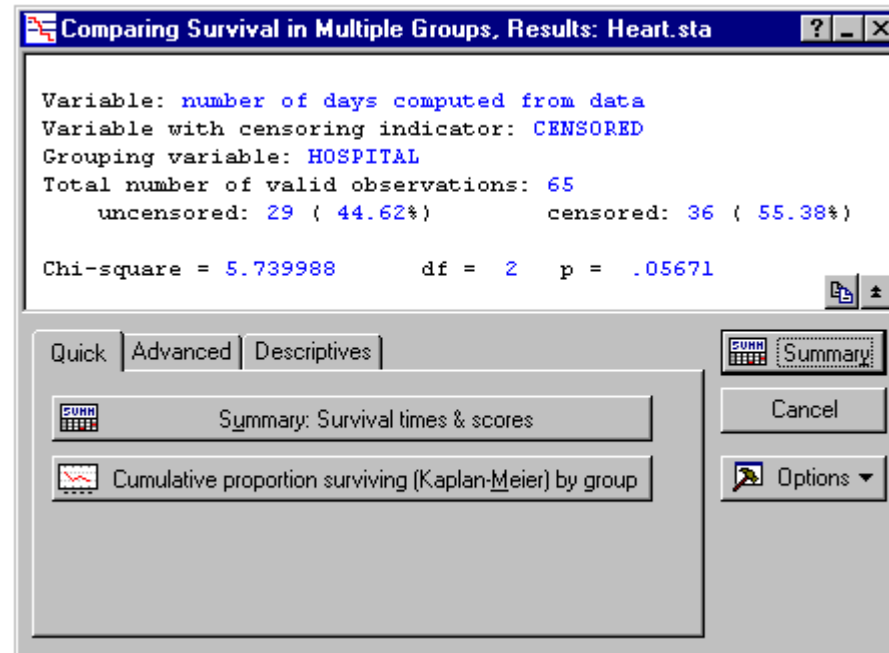
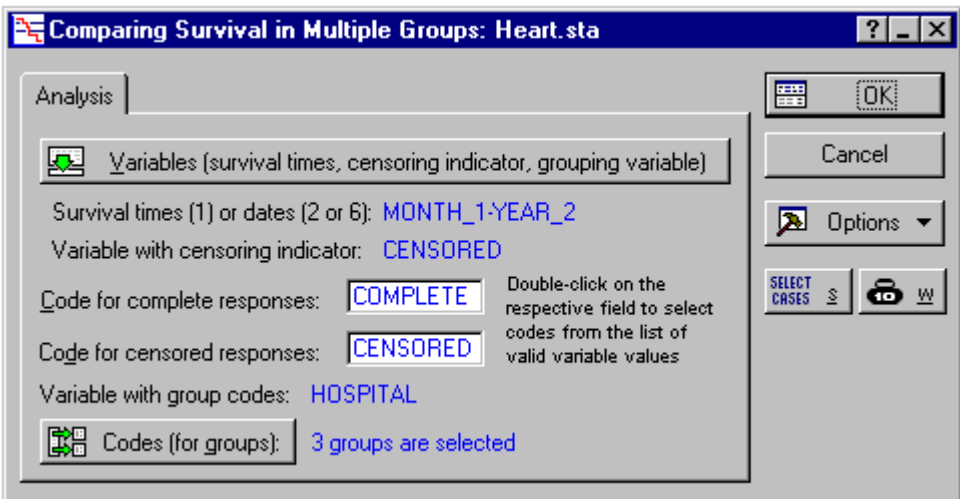
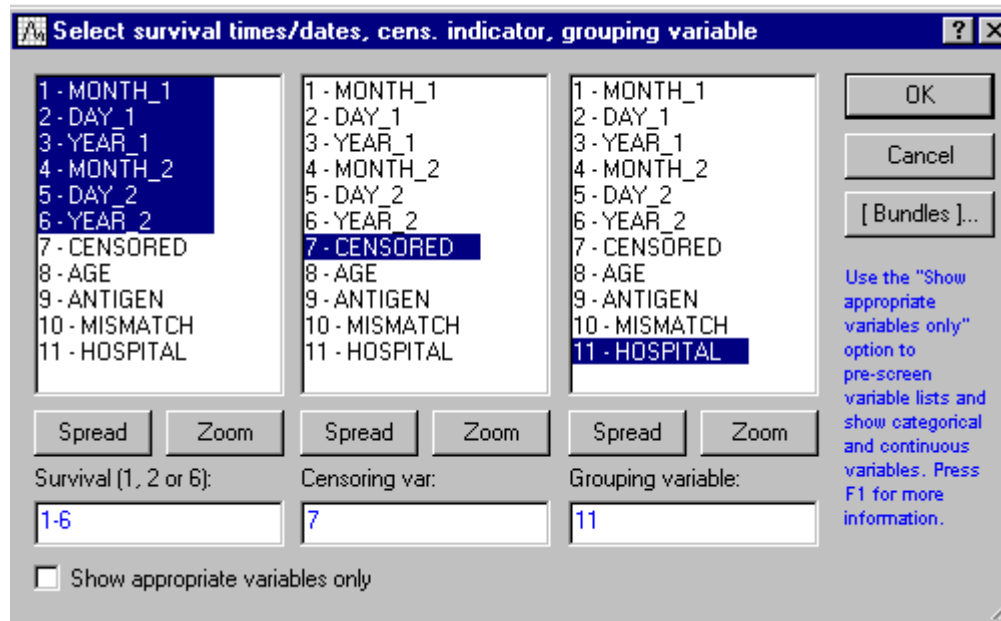
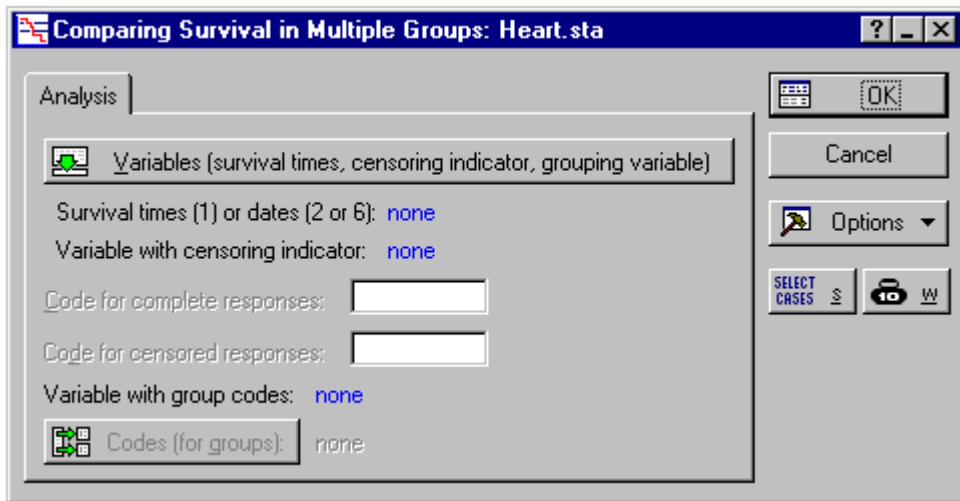
Life Table (Heart.sta)

Log-Likelihood for data: -68.6809

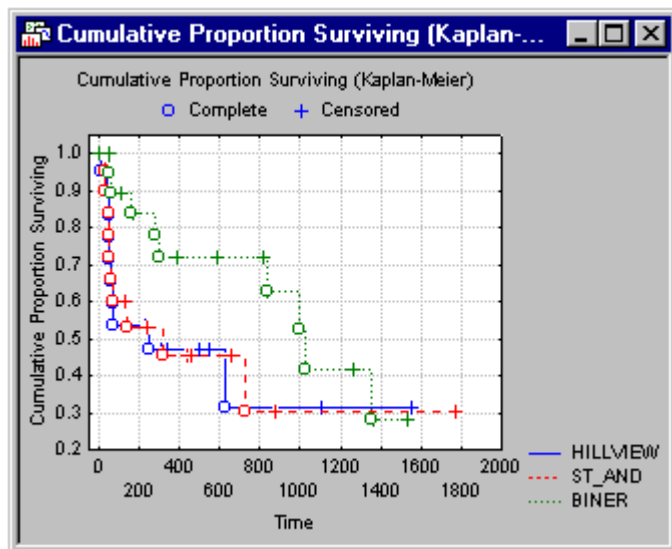
Interval	Interval Start	Mid Point	Interval Width	Number Entering	Number Withdrawn	Number Exposed	Number Dying	Proportn Dead	Proportn Surviving
Intno.1	0.000	80.682	161.3636	65	14	58.00000	19	0.327586	0.672414
Intno.2	161.364	242.046	161.3636	32	4	30.00000	4	0.133333	0.866667
Intno.3	322.727	403.409	161.3636	24	4	22.00000	0	0.022727	0.977273
Intno.4	484.091	564.773	161.3636	20	4	18.00000	1	0.055556	0.944444
Intno.5	645.455	726.136	161.3636	15	1	14.50000	1	0.068966	0.931034
Intno.6	806.818	887.500	161.3636	13	3	11.50000	1	0.086957	0.913044
Intno.7	968.182	1048.864	161.3636	9	1	8.50000	2	0.235294	0.764706
Intno.8	1129.545	1210.227	161.3636	6	1	5.50000	0	0.090909	0.909091
Intno.9	1290.909	1371.591	161.3636	5	1	4.50000	1	0.222222	0.777778
Intno.10	1452.273	1532.955	161.3636	3	2	2.00000	0	0.250000	0.750000
Intno.11	1613.636	1694.318	161.3636	1	0	1.00000	0	0.500000	0.500000
Intno.12	1775.000			1	1	0.50000	0	1.000000	0.000000



Túlélési görbék összehasonlítása



Túlélési görbék összehasonlítása



Comparing Survival in Two Groups: Heart.sta

Analysis

Variables (survival times, censoring indicator, grouping variable)

Survival times (1) or dates (6): MONTH_1-YEAR_2

Variable with censoring indicator: CENSORED

Code for complete responses: COMPLETE

Code for censored responses: CENSORED

Variable with group codes: HOSPITAL

Code for first group: HILLVIEW

Code for second group: BINER

Buttons: OK, Cancel, Options, SELECT CASES, 10, W

Two-Sample Tests Results: Heart.sta

Variable: number of days computed from data

Variable with censoring indicator: CENSORED

Grouping variable: HOSPITAL

Total number of valid observations: 43

uncensored: 19 (44.19%) censored: 24 (55.81%)

Valid observations: Group 1 (HILLVIEW): 22 Group 2 (BINER): 21

 Uncensored: 10 (45.45%) 9 (42.86%)

 Censored: 12 (54.55%) 12 (57.14%)

Buttons: Quick, Two-sample tests, Function plots, Gehan test, Cancel, Options

Cox Regression Modellek

Select: times/dates, indep. vars., cens. var., grouping var. (optional)

1 - MONTH_1 2 - DAY_1 3 - YEAR_1 4 - MONTH_2 5 - DAY_2 6 - YEAR_2 7 - CENSORED 8 - AGE 9 - ANTIGEN 10 - MISMATCH 11 - HOSPITAL	1 - MONTH_1 2 - DAY_1 3 - YEAR_1 4 - MONTH_2 5 - DAY_2 6 - YEAR_2 7 - CENSORED 8 - AGE 9 - ANTIGEN 10 - MISMATCH 11 - HOSPITAL	1 - MONTH_1 2 - DAY_1 3 - YEAR_1 4 - MONTH_2 5 - DAY_2 6 - YEAR_2 7 - CENSORED 8 - AGE 9 - ANTIGEN 10 - MISMATCH 11 - HOSPITAL	1 - MONTH_1 2 - DAY_1 3 - YEAR_1 4 - MONTH_2 5 - DAY_2 6 - YEAR_2 7 - CENSORED 8 - AGE 9 - ANTIGEN 10 - MISMATCH 11 - HOSPITAL
--	--	--	--

Survival (1, 2 or 6): Indep. variables: Censuring variable: Grouping (optional):

Show appropriate variables only

Regression Models for Censored Data: Heart.sta

Model: **Proportional hazard (Cox) regression**

Variables (survival times, indep., censoring, [optional] grouping):

Survival times (dependent variable): **MONTH_1-YEAR_2**

Independent variables: **AGE-MISMATCH**

Variable with censoring indicator: **CENSORED**

Code for complete responses: **COMPLETE**

Code for censored responses: **CENSORED**

Variable with group codes (for stratified analysis): **none**

Codes (for groups): 3 groups are selected

MD deletion: Casewise Mean substitution

To specify the dependent variable, select either 1 variable (e.g., survival times), or 2 or 6 variables (start- and end-times or dates).

Regression Results: Heart

Model: **Proportional hazard (Cox) regression**

Variables
 dependent: **number of days computed from data**
 independent: **AGE ANTIGEN MISMATCH**
 Variable with censoring indicator: **CENSORED**

Total number of valid observations: **65**
 uncensored: **29 (44.62%)** censored: **36 (55.38%)**

Log-Likelihood of final solution: **-87.8674**
 Log-Likelihood of Null model (all B's=0): **-99.0270**
 Chi-Square (Null model - final solution): **22.31933 df= 3 p= .00006**

Conf. Level:

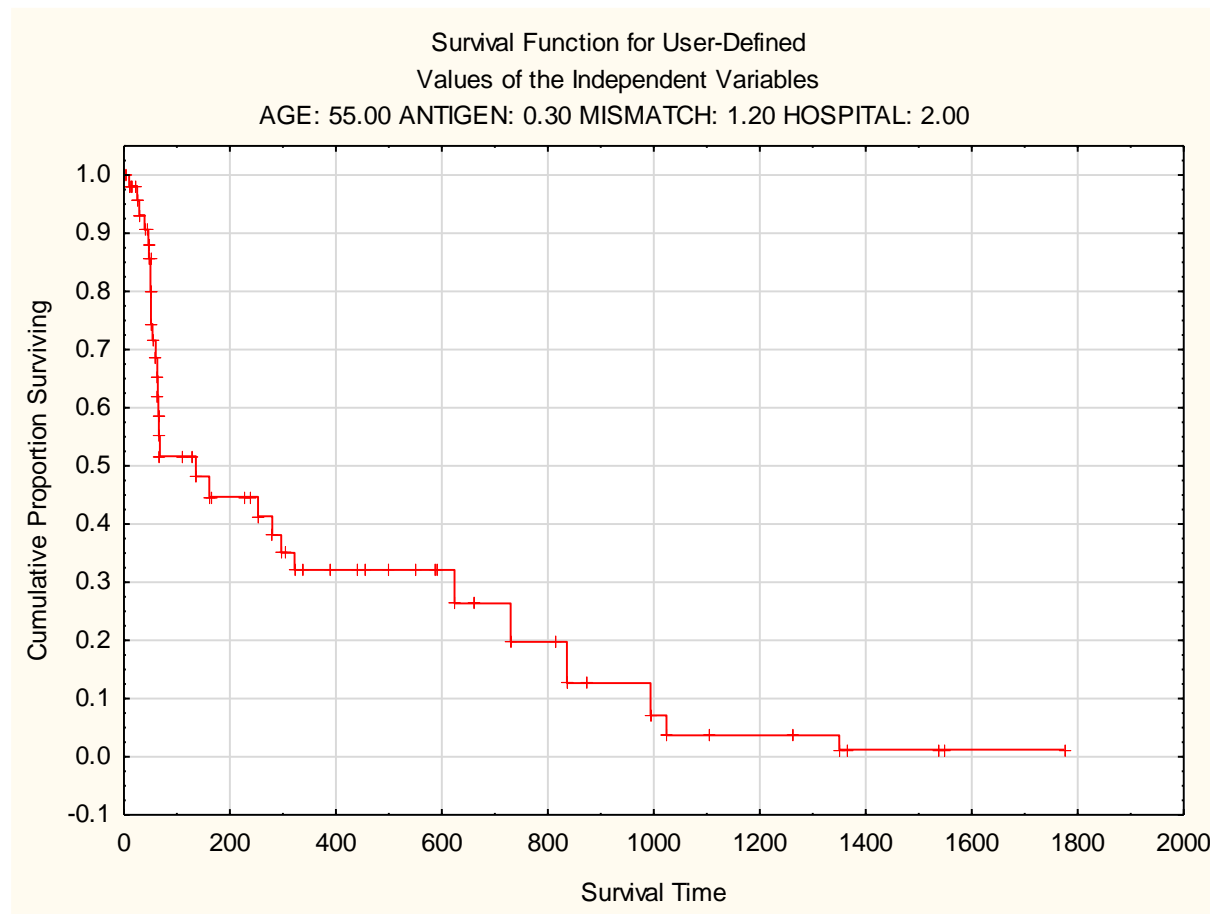
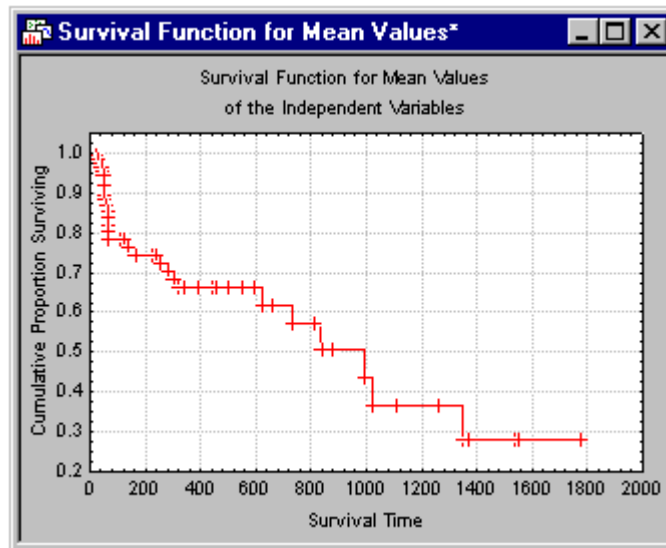
Summary: Parameter estimates

Stratified analysis

Data: Dependent Variable: Survival times in days (Heart)*

Dependent Variable: Survival times in days (Heart)
 Censuring var.: CENSORED
 Chi² = 22.3193 df = 3 p = .00006

	Beta	Standard Error	Beta 95% lower	Beta 95% upper	t-value	Wald Statist.	p	Risk ratio	Risk ratio 95% lower	Risk ratio 95% upper
N=65										
AGE	0.109096	0.033293	0.043843	0.174349	3.276836	10.73766	0.001051	1.115269	1.044818	1.190470
ANTIGEN	-0.048782	0.471644	-0.973188	0.875623	-0.103431	0.01070	0.917622	0.952388	0.377876	2.400371
MISMATCH	1.063761	0.394599	0.290362	1.837160	2.695804	7.26736	0.007026	2.897246	1.336911	6.278683



File Edit View Insert Format Statistics Data Mining Graphs Tools Data Window Help

Arial 10 B I U [Text Alignment Icons] [Color/Font Icons] [Language/Encoding Icons] Vars Cases

Heart transplant data		Resume... Ctrl+R						
	1 MONTH_1	2 DAY_1	7 CENSORED	8 AGE	9 ANTIGEN	10 MISMATCH	11 HOSPITAL	
1	JANUARY	6	CENSORED	54	0	1.11	HILLVIEW	
2	MAY	2	CENSORED	40	0	1.66	HILLVIEW	
3	AUGUST	31	COMPLETE	51	0	1.32	HILLVIEW	
4	AUGUST	22	COMPLETE	42	0	0.61	ST_AND	
5	SEPTEMBER	9	CENSORED	48	0	0.36	ST_AND	
6	OCTOBER	5	COMPLETE	54	0	1.89	ST_AND	
7	OCTOBER	26				0.87	BINER	
8	NOVEMBER	22				1.12	BINER	
9	NOVEMBER	20				2.05	HILLVIEW	
10	FEBRUARY	15				2.76	HILLVIEW	
11	FEBRUARY	8				1.13	BINER	
12	MARCH	29				1.38	HILLVIEW	
13	APRIL	13				0.96	ST_AND	
14	JULY	16				1.62	ST_AND	
15	MAY	22				1.06	ST_AND	
16	AUGUST	16				0.47	BINER	
17	SEPTEMBER	3				1.58	BINER	
18	SEPTEMBER	14				0.69	HILLVIEW	
19	JANUARY	16				0.91	BINER	
20	JANUARY	3	70	APRIL	1	74	0.38	HILLVIEW
21	MAY	19	70	JULY	12	70	2.09	HILLVIEW
22	MAY	13	70	JUNE	29	70	0.87	ST_AND
23	MAY	9	70	MAY	9	70	0.87	ST_AND
24	JULY	4	70	APRIL	1	74	0.75	BINER
25	OCTOBER	15	70	APRIL	1	74	0.98	BINER
26	JANUARY	5	71	FEBRUARY	18	71	0.00	ST_AND
27	JANUARY	11	71	OCTOBER	1	73	0.81	BINER

- Basic Statistics/Tables
- Multiple Regression
- ANOVA
- Nonparametrics
- Distribution Fitting
- Distributions & Simulation
- Advanced Linear/Nonlinear Models
- Multivariate Exploratory Techniques
- Industrial Statistics & Six Sigma
- Power Analysis
- Automated Neural Networks
- PLS, PCA, Multivariate/Batch SPC
- Variance Estimation and Precision
- Statistics of Block Data
- Statistica Visual Basic
- Batch (ByGroup) Analysis
- Probability Calculator

- General Linear Models
- Generalized Linear/Nonlinear Models
- Stepwise Model Builder
- General Regression Models
- General Partial Least Squares Models
- NIPALS Algorithm (PCA/PLS)
- Variance Components
- Survival Analysis
- Cox Proportional Hazards Models
- Nonlinear Estimation
- Fixed Nonlinear Regression
- Log-Linear Analysis of Frequency Tables
- Time Series/Forecasting
- Structural Equation Modeling

File Edit View Insert Format Statistics Data Mining Graphs Tools Data Window Help

Arial 10 B I U [Text Alignment Icons] [Number Formatting Icons] [Language Icons] Vars Cases [Grid Icons]

Heart transplant data from Crowley and Hu, stratified

	1	2	3	4	5	6	7	8	9	10	11
	MONTH_1	DAY_1	YEAR_1	MONTH_2	DAY_2	YEAR_2	CENSORED	AGE	ANTIGEN	MISMATCH	HOSPITAL
1	JANUARY	6	68	JANUARY	21	68	CENSORED	54	0	1.11	HILLVIEW
2	MAY	2	68	MAY	5	68	CENSORED	40	0	1.66	HILLVIEW
3	AUGUST	31	68	MAY	17	70	COMPLETE	51	0	1.32	HILLVIEW
4	AUGUST	22	68	OCTOBER	7	68	COMPLETE	42	0	0.61	ST_AND
5	SEPTEMBR	9	68	JANUARY	14	69	CENSORED	48	0	0.36	ST_AND
6	OCTOBER	5	68	DECEMBER	8	68	COM				
7	OCTOBER	26	68	JULY	7	72	COM				
8	NOVEMBER	22	68	AUGUST	29	69	COM				
9	NOVEMBER	20	68	DECEMBER	13	68	CEN				
10	FEBRUARY	15	69	FEBRUARY	25	69	COM				
11	FEBRUARY	8	69	NOVEMBER	29	71	COM				
12	MARCH	29	69	MAY	7	69	COM				
13	APRIL	13	69	APRIL	13	71	COM				
14	JULY	16	69	NOVEMBER	29	69	COM				
15	MAY	22	69	APRIL	1	74	CEN				
16	AUGUST	16	69	AUGUST	17	69	CEN				
17	SEPTEMBR	3	69	DECEMBER	18	71	COM				
18	SEPTEMBR	14	69	NOVEMBER	13	69	COM				
19	JANUARY	16	70	APRIL	1	74	CEN				
20	JANUARY	3	70	APRIL	1	74	CEN				
21	MAY	19	70	JULY	12	70	COM				
22	MAY	13	70	JUNE	29	70	COM				
23	MAY	9	70	MAY	9	70	CEN				
24	JULY	4	70	APRIL	1	74	CEN				
25	OCTOBER	15	70	APRIL	1	74	CEN				
26	JANUARY	5	71	FEBRUARY	18	71	CEN				
27	JANUARY	11	71	OCTOBER	1	73	COM				
28	FEBRUARY	22	71	APRIL	14	71	COM				
29	MARCH	22	71	APRIL	1	74	CENSORED	36	0	1.35	HILLVIEW
30	APRIL	24	71	JANUARY	2	72	COMPLETE	48	1	1.08	HILLVIEW
31	AUGUST	18	71	OCTOBER	8	71	COMPLETE	52	0	1.51	ST_AND
32	NOVEMBER	8	71	APRIL	1	74	CENSORED	38	0	0.98	ST_AND
33	OCTOBER	13	71	AUGUST	30	72	COMPLETE	48	1	1.82	ST_AND
34	DECEMBER	15	71	APRIL	1	74	CENSORED	41	0	0.19	BINER
35	NOVEMBER	20	71	JANUARY	9	72	COMPLETE	49	0	0.66	BINER
36	JANUARY	7	72	APRIL	1	74	CENSORED	32	1	1.93	BINER
37	MARCH	4	72	SEPTEMBR	6	73	CENSORED	48	0	0.12	HILLVIEW
38	MARCH	17	72	MAY	22	72	COMPLETE	51	0	1.12	HILLVIEW
39	MAY	18	72	JANUARY	1	73	CENSORED	19	0	1.02	HILLVIEW

Cox Proportional Hazards Results: Heart

Quick | Survivor Function | Plots

Goodness of Fit
Iteration history
Parameter estimates
Type 3 Tests
Design Terms
Covariance matrix
Correlation matrix
Estimate

Residuals spreadsheet
Residuals Save
 Martingale Dfbeta
 Deviance Xbeta
 Score StdXBeta
 Schoenfeld Likelihood displacement
 Sc. Schoenfeld Lmax

Test of proportionality
 PH Test
 Sc. Schoen Plots

Transformation
 Identity
 Log
 KM
 Rank

Options
Modify
Code generator
By Group...

Conf. level: .950

Cox *proportional* hazards (PH) model

Cox Proportional Hazards Results: Heart

Quick | Survivor Function | Plots

Number of covariate patterns: 1

Plot

Survivor function

Survivor function

Estimator

Breslow

Kaplan-Meier

Add confidence interval

Type

Log

Normal

Log-Log

Summary

Cumulative Hazard function

Cumulative Hazard

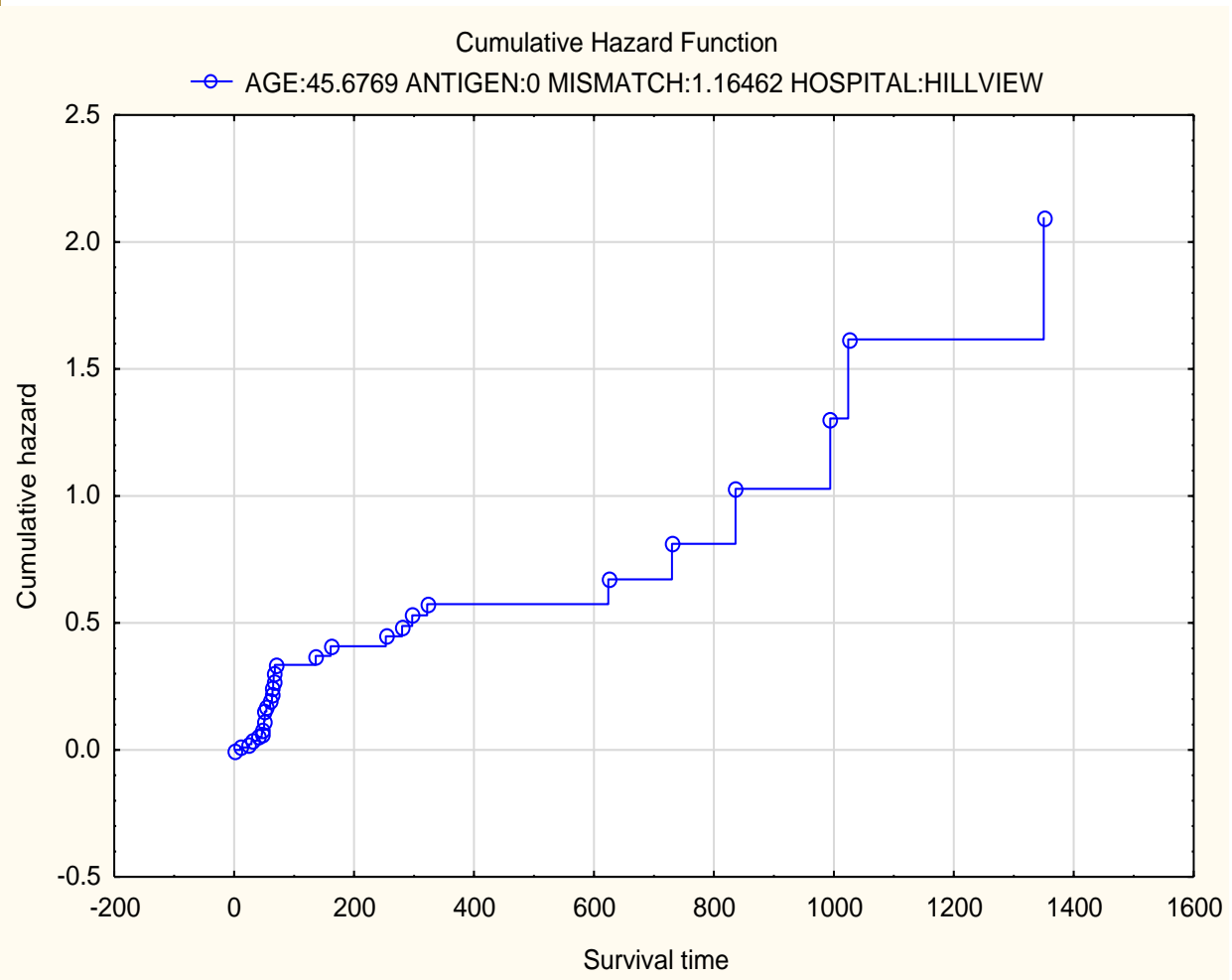
Add confidence interval

Options

Modify

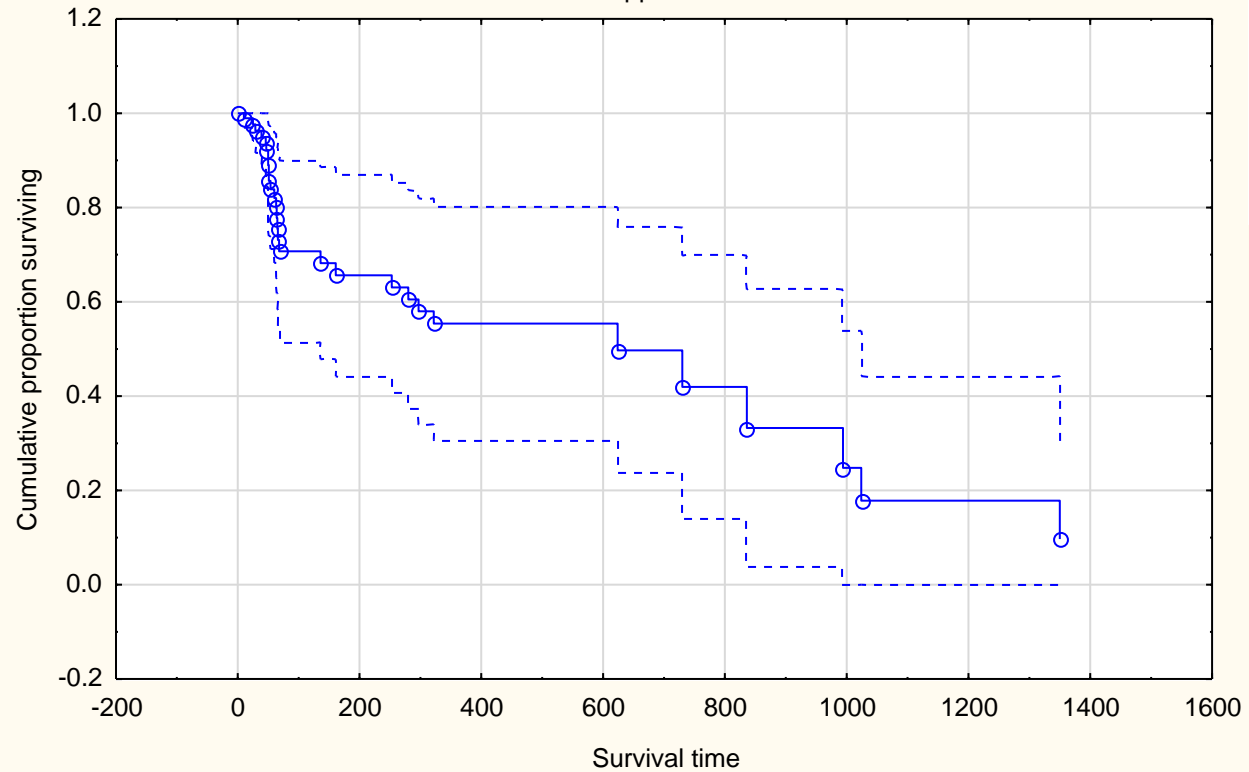
Code generator

By Group...



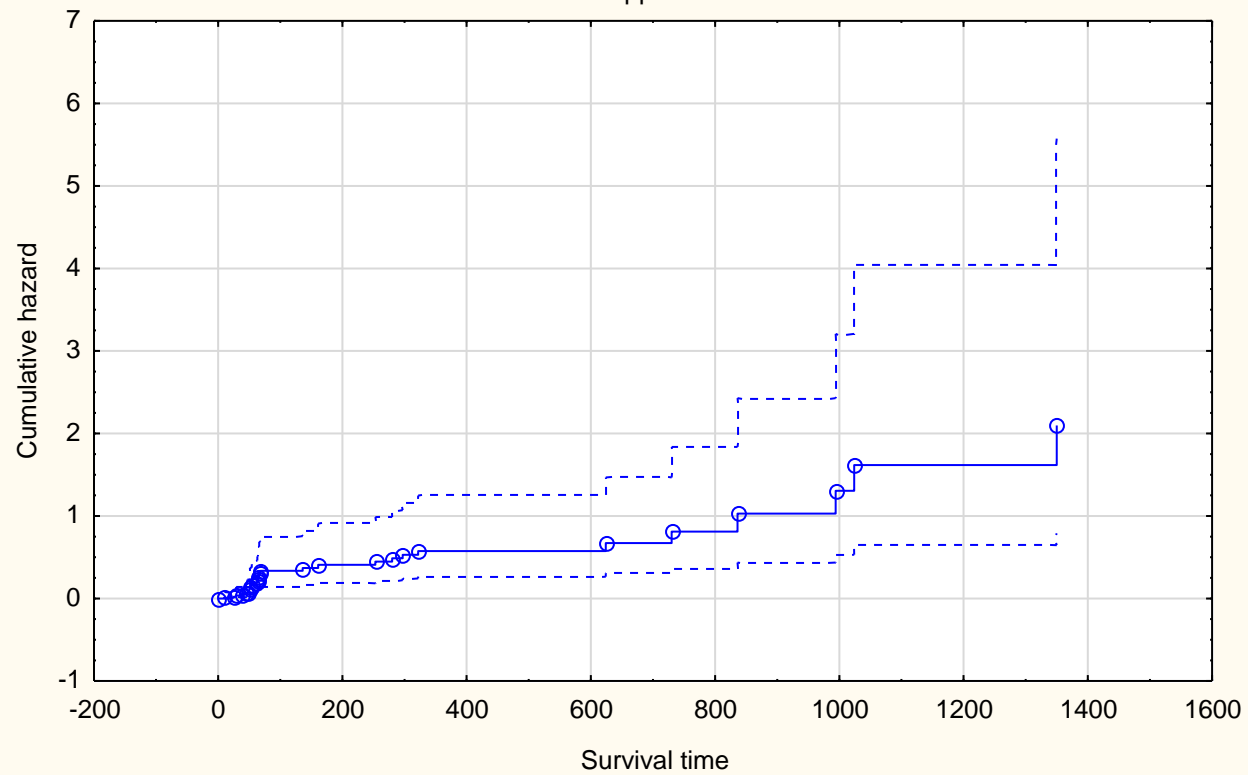
Survivor Function: Kaplan-Meier

—○ AGE:45.6769 ANTIGEN:0 MISMATCH:1.16462 HOSPITAL:HILLVIEW
- - - 95%LowerNormal
- - - 95%UpperNormal



Cumulative Hazard Function

—○ AGE:45.6769 ANTIGEN:0 MISMATCH:1.16462 HOSPITAL:HILLVIEW
- - - 95%LowerNormal
- - - 95%UpperNormal



Köszönöm a figyelmet!