Déjà

1 (homogeneous) sample: "Survival" / "Time-to-event" data:

- [equivalent] Functions: S[t] , hazard h[t] , pdf[t]
- Links: e.g. $S[t] = \exp[-\int h[u]\, du]$, integral from u=0 to u=t
- Summaries of these functions (e.g. T_{25}, T_{50}, S[T])
- ***Non-Parametric / Semi-Parametric*** Estimation

  (point & interval) of S[t] , h[t] and pdf[t]

  --- Lifetable [fixed-] --- K-M [data-determined intervals]
- ***Censored data not necessarily " time - to - event "***

  Y = PSA levels < detection limit, salaries in intervals, distance travelled on set of tires, pages on single ink cartridge, etc.
- ***Not covered, but possible: Parametric models****

Comparison of 2 Survival/Hazard Curves or Distributions

- Risk Sets
- Adjusted comparisons (non-regression methods)
- ***Not covered, but possible: Parametric models****

---

* The **SAS LIFEREG** procedure fits parametric models to failure time data that can be right, left, or interval censored. The models for the response variable consist of a linear effect composed of the covariates and a random disturbance term. The distribution of the random disturbance can be taken from a class of distributions that includes the extreme value, normal, logistic, and, by using a log transformation, the exponential, Weibull, lognormal, loglogistic, and gamma distributions.

Stata streg performs maximum likelihood estimation of parametric regression survival-time models. Survival models currently supported are exponential, Weibull, Gompertz, lognormal, log-logistic and generalized gamma. Also see help stcox for estimation of proportional hazards models.

New

Regression Models (semiparametric)

- Model (event) ***rates*** or ***hazards*** p ***h*** reg
- Models are multiplicative in rates/hazards

  linear predictors if work in log[rate] or log[hazard] scale
- Proportionality of rates or hazards ***p h*** reg

  Constancy of rate ratio parameter over time-bands
- Avoid modelling the nuisance parts

  don't fit parameters that (a) are not our focus (b) waste "d.f."
- Use risksets & conditioning to reduce # parameters
- Choice of Time-scale and "Time-zero" is important

  (has implications for risksets)
- Models, and conditioning as a way of eliminating parameters, applicable to matched case-control studies and even to c-c and other (e.g. consumer choice*) studies with no 'time' element

  (* Daniel McFadden shared the Nobel Prize for his development of theory and methods for analyzing discrete choice in Economics: http://www.nobel.se/economics/laureates/2000/mcfadden-autobio.html)

Readings (* = 3 most relevant for now, to be handed out in class 08/03)

[http://www.epi.mcgill.ca/hanley/c681/cox]

* Clayton&Hills, Ch 22 Intro to regression models

* Pair of expository articles by JH

* Clayton&Hills, Ch 30 Cox's Method, specifically... pp 298; bottom of pp 300 to top of p 304; Exercise 30.1

Data: Framingham Study

[http://www.epi.mcgill.ca/hanley/c681/cox]

Other Resources

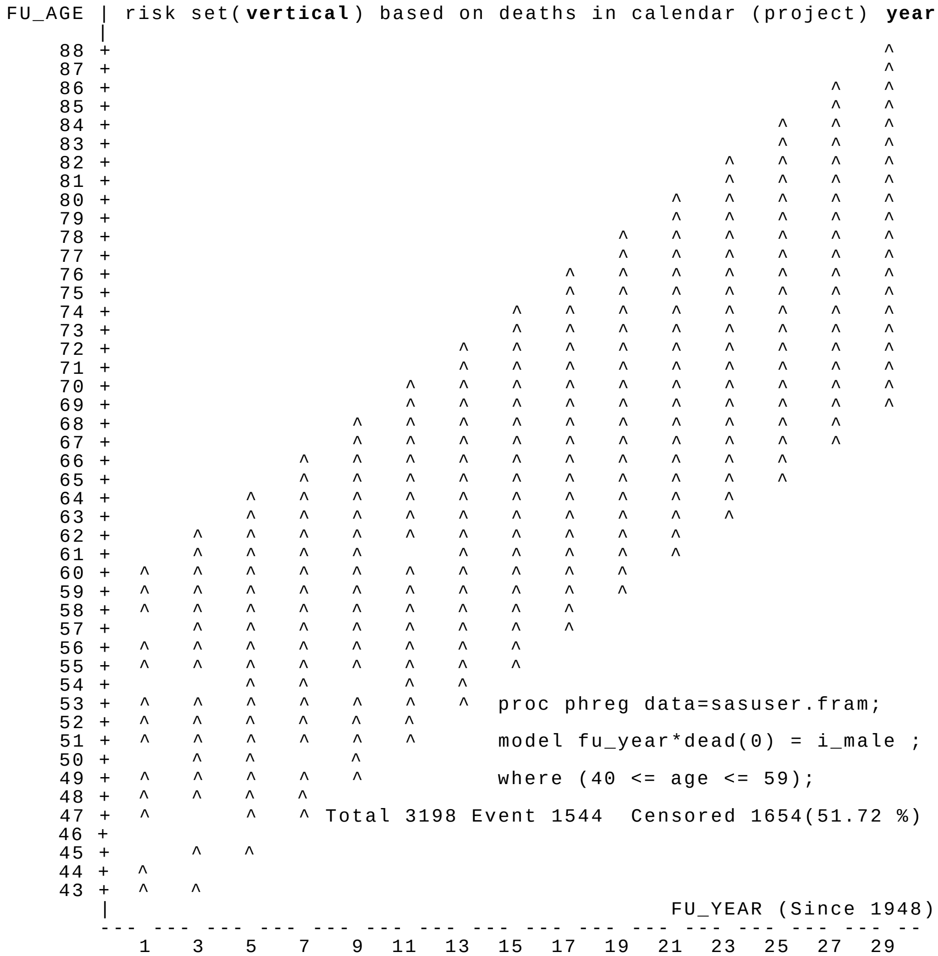
- Texts [http://www.epi.mcgill.ca/hanley/c681/cox]

  Kleinbaum's 'Self-Learning' textbook, Chapter 3

  Collett Textbook, Chapter 3

  Hosmer & Lemeshow's Applied Survival Analysis

.... **TIME Scale = YEAR_of_research_grant** ...



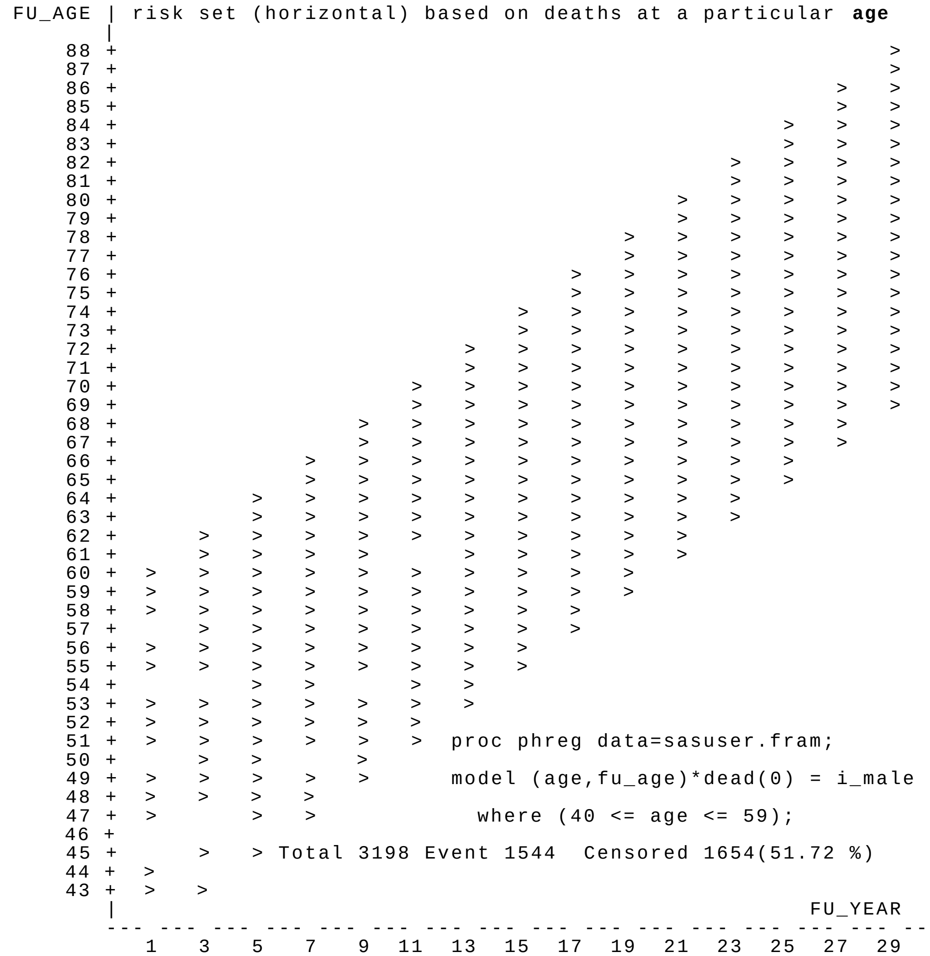
Testing Global Null Hypothesis: BETA=0

-2LOGL W/out:24098.1 With:23968.1 Covariates; Chi-Sq(1) 130; p=0.0001

Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Sq	Pr > Chi-Sq	Risk Ratio
I_MALE	1	0.583	0.051	129.3	0.0001	1.79

TIME Scale = AGE .(NOTE how delayed entry is specified)



Testing Global Null Hypothesis: BETA=0

-2LOGL W/out:22819.8 With:22662.7 Covariates; Chi-Sq(1) 157; p=0.0001

Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Sq	Pr > Chi-Sq	Risk Ratio
I_MALE	1	0.643	0.051	156.3	0.0001	1.90