Models In Epidemiology And Biostatistics
Gordon Hilton Fick

Classical Methods With Time-To-Event Studies

Continuous and Discrete

We now direct our attention to studies in which one is interested in the time to an event. The event might be death [in which case we speak of survival studies]. The event might be a recurrence of a form of cancer. The event can be positive in nature like, for example, in fertility studies, the time to pregnancy. Some authors refer to survival studies no matter the nature of the event of interest.

Time-to-event studies usually fall into two quite separate types: Absolutely Continuous and Discrete.

If the investigator is able to measure the time to event with some accuracy and definiteness, we can typically use methods that assume continuity of time. For example, mortality is usually known to the day. With 'real' continuous time outcomes, individuals with identical times [ties] are very rare.

If the investigator can only determine the time to event within a finite set of intervals of time, then we typically use methods that acknowledge the discreteness of time. For example, a participant may be observed only at specific visits to a clinic. The participant may then be noted as having experienced the event during the time between the last visit and the current visit.

Inevitably, there are studies in which time is essentially discrete but, for various reasons, the analyst makes an assumption of approximate continuity. Sometimes, one of challenges with such a plan can be that the resulting outcomes contain a number of 'ties' that cannot be ignored and then methods must be used to correctly address the ties issue.

There are also circumstances in which continuous time outcomes are grouped into intervals and then analyzed using discrete methods.

Many of the 'classic' discrete methods were developed a long time ago [like a 100 years ago and more] while the classic continuous methods were apparently first developed in the 1950's. 'Modern' model based methods [for either continuous or discrete time outcomes] began to be developed in the 1970's and really came into their own with the pioneering work of DR Cox and R Prentice [to name just two].

Censoring

[Wikipedia] Censoring is a condition in which the value of a measurement or observation is only partially known.
For example, suppose a study is conducted to measure the impact of a drug on mortality rate. In such a study, it may be known that an individual's age at death is at least 75 years (but may be more). Such a situation could occur if the individual withdrew from the study at age 75, or if the individual is currently alive at the age of 75.
Censoring also occurs when a value occurs outside the range of a measuring instrument. For example, a bathroom scale might only measure up to 300 pounds (140 kg). If a 350 lb (160 kg) individual is weighed using the scale, the observer would only know that the individual's weight is at least 300 pounds (140 kg).
The problem of censored data, in which the observed value of some variable is partially known, is related to the problem of missing data, where the observed value of some variable is unknown.
Censoring should not be confused with the related idea truncation. With censoring, observations result

either in knowing the exact value that applies, or in knowing that the value lies within an interval. With truncation, observations never result in values outside a given range: values in the population outside the range are never seen or never recorded if they are seen. Note that in statistics, truncation is not the same as rounding.
There are many forms of censoring. A participant may be followed until the end of the study but has not experienced the event and the study has ended at time t, say. Then we only know that the time to event is at least the time in which the study ended for that person. [that time-to-event is greater than t]. A participant might be followed until a time when it is recognized that the participant is 'loss to follow up'. Then the investigator attempts to assign a time t for this person and again one then only knows that the time-to-event is greater than t. Time to 'loss to follow up' is rarely known with precision. We may have only an interval of time here.

Forms of censoring:
Left censoring – a data point is below a certain value but it is unknown by how much.
Interval censoring – a data point is somewhere on an interval between two values.
Right censoring – a data point is above a certain value but it is unknown by how much.
Type I censoring occurs if an experiment has a set number of subjects or items and stops the experiment at a predetermined time, at which point any subjects remaining are right-censored.
Type II censoring occurs if an experiment has a set number of subjects or items and stops the experiment when a predetermined number are observed to have failed; the remaining subjects are then right-censored.
Random (or non-informative) censoring is when each subject has a censoring time that is statistically independent of their failure time. The observed value is the minimum of the censoring and failure times; subjects whose failure time is greater than their censoring time are right-censored.

Interval censoring can occur when observing a value requires follow-ups or inspections. Left and right censoring are special cases of interval censoring, with the beginning of the interval at zero or the end at infinity, respectively.
A common misconception with time interval data is to class as left censored intervals where the start time is unknown. In these cases we have a lower bound on the time interval, thus the data is right censored (despite that fact that the missing start point is to the left of the known interval when viewed as a timeline).
One of the earliest attempts to analyze a statistical problem involving censored data was Daniel Bernoulli's 1766 analysis of smallpox morbidity and mortality data to demonstrate the efficacy of vaccination.

Competing Events

There are many forms of competing events. Suppose there is a primary time outcome; say, for example, death due to cancer. A competing event could be death from cardiovascular disease. This competing event occurring precludes the occurrence of the primary event. Sometimes, a competing event occurrence can fundamentally alter the probability of occurrence of the primary event.
For example, [from Pintilie 2006] in cancer research, patients undergo one or more of the three main types of treatment: surgery, chemotherapy and radiation therapy. Suppose that the treatment appeared successful and all evidence of disease was removed. A common endpoint of interest in cancer studies is the time to the return of disease (relapse) after the initial apparent success of treatment. The relapse may be at the site of the initial disease, in which case the endpoint is called local relapse, or at a different site, called distant relapse or metastasis.
Chemotherapy, as a systemic treatment, affects the whole body while both surgery and radiation therapies are treatments directed towards the specific disease site. Therefore, in studies of radiation or surgery, the researcher may be more interested in the time to local relapse than in the time to metastasis

or death. In this case, it is desirable to identify characteristics that are associated with local relapse. However, a patient may develop distant disease and die before a local relapse is observed. In this case the observation of distant disease hinders the observation of local disease. Furthermore, the occurrence of local disease after distant disease may not be of much interest since treatment of the distant disease may alter the chances of local disease recurring.
More generally, the term 'relapse' refers to the return or recurrence of any potentially chronic condition or disease after an initial improvement. For instance, in studies of smoking cessation, relapse refers to the resumption of the previous smoking behaviour. Psychiatric studies of patients with bipolar disorder might consider a repeat episode of mania to be a relapse, while a clinical trial involving patients with chronic bronchitis might define a relapse to be the reappearance of the symptoms. Pulmonary tuberculosis could recur either within the lung (local relapse) or at extrapulmonary sites (other relapse). Similarly, a relapse of the herpes simplex virus could be experienced as a skin lesion (local relapse) or within the central nervous system (other relapse). Therefore, the concept of relapse (local, distant or other) is a meaningful endpoint in other medical areas, beside cancer, where the disease or condition under study may reappear at different sites within the body.

Introduction to Discrete Time Methods
Let us suppose we have a collection of time intervals [mutually exclusive and exhaustive]. We will label the intervals by the right hand endpoint of each interval. Say then we have $t_1, t_2, \ldots t_k$ where t_i indicates the right hand endpoint of the ith interval. An event occurring in the interval labeled t_i has probability $p(t_i) = p_i$. The distribution function [also called the failure function] is $F(t) = \sum_{t_j \le t} p_j$ and the survivor function is $S(t) = \sum_{t_j > t} p_j$. We will also use the notation $S(t_i) = P_i$. We can compute estimates [with standard errors and confidence intervals] for the p_i and P_i using methods called 'life table' methods.

An example should help here.

```
. use selvin.dta
. ltable time died
```

Interval		Beg. Total	Deaths	Lost	Survival	Std. Error	[95% Conf.	Int.]
0	1	40	2	9	0.9437	0.0387	0.7930	0.9856
1	2	29	2	6	0.8711	0.0609	0.6890	0.9501
2	3	21	4	1	0.7011	0.0906	0.4844	0.8403
3	4	16	3	3	0.5561	0.1036	0.3351	0.7298
4	5	10	2	1	0.4390	0.1100	0.2243	0.6354
5	6	7	2	1	0.3039	0.1101	0.1152	0.5188
7	8	4	1	3	0.1823	0.1150	0.0296	0.4391

We can see that, for example, 0.9437 is an estimate of P_1 . Hence, it is an estimate of the probability of survival to the end of the first interval or the probability that the time to the event is greater than $t_1 = 1$.
One can get a graph of the estimates and confidence intervals using:

```
. ltable time died, graph ci
```

Using the default to connect the estimates gives lines drawn between estimates [suggesting a linear decline in survival probability between estimates.
Alternately, one can use:

```
. ltable time died, graph ci plotopts(connect(none))
```

which acknowledges that the estimates apply to the intervals and the inherent discreteness of time.

We can compare life tables. For example:

```
. use pike.dta
. ltable t died, by(group) interval(30)

                Beg.                                    Std.
   Interval     Total   Deaths   Lost    Survival     Error      [95% Conf. Int.]
-------------------------------------------------------------------------------
group = 1
  120   150        19        1      0      0.9474    0.0512      0.6812    0.9924
  150   180        18        1      0      0.8947    0.0704      0.6408    0.9726
  180   210        17        6      0      0.5789    0.1133      0.3321    0.7626
  210   240        11        6      1      0.2481    0.1009      0.0847    0.4552
  240   270         4        2      1      0.1063    0.0786      0.0139    0.3090
  300   330         1        1      0      0.0000         .           .         .
group = 2
  120   150        21        1      0      0.9524    0.0465      0.7072    0.9932
  150   180        20        2      0      0.8571    0.0764      0.6197    0.9516
  180   210        18        2      1      0.7592    0.0939      0.5146    0.8920
  210   240        15        7      0      0.4049    0.1099      0.1963    0.6053
  240   270         8        2      0      0.3037    0.1031      0.1245    0.5057
  270   300         6        4      0      0.1012    0.0678      0.0172    0.2749
  300   330         2        1      0      0.0506    0.0493      0.0035    0.2073
  330   360         1        0      1      0.0506    0.0493      0.0035    0.2073
-------------------------------------------------------------------------------
```

For example, to compare group 1 with group 2 at the fourth time interval, one can compute a z test:
$z = (0.2481 - 0.4049)/\sqrt{(0.1009)^2 + (0.1099)^2} = -1.0509813$ p-value = 0.2932

Time interval specific comparisons can be vastly more informative than any single omnibus test attempting to compare two survivor functions. For example, it might happen one that time specific comparison indicates group1 survival higher than group 2 survival while a different time interval comparison indicates the reverse. Maybe, then, early time comparisons suggest group 1 better off than group 2 but at a later time group 2 ends up being better off than group 1.

Introduction to Continuous Time Methods

Now, the outcome t [= time until event] has a density function: f(t). The area under this curve gives us probability. We continue to consider the distribution function F(t)= P((0,t]) and the survival function S(t)=1-F(t). So F gives us probability up to time t while S gives probability beyond t. The term survival function works well in contexts where the event is death and so S(t) is indeed the probability for survival to time t. In some fields the study of Survival is primary while in other fields Failure [F(t)] is primary. Of course, knowing one determines the other.
Now we are interested in constructing an estimate of the [entire] function S(t) which we will write:
$\hat{S}(t)$. We will be interested in comparing survival functions as well.

Perhaps the most direct method of analysis is based on the premise that the logarithm of time t : log(t) may be approximately symmetrically distributed and so one can try 'simple' regression analysis.
To illustrate, we will consider a study with no censoring or competing events so that least squares regression can be tried out: $\log(t) = \beta_0 + \beta_1 G + \epsilon$

```
use intro_surv_1.dta
gen lt=log(t)
```

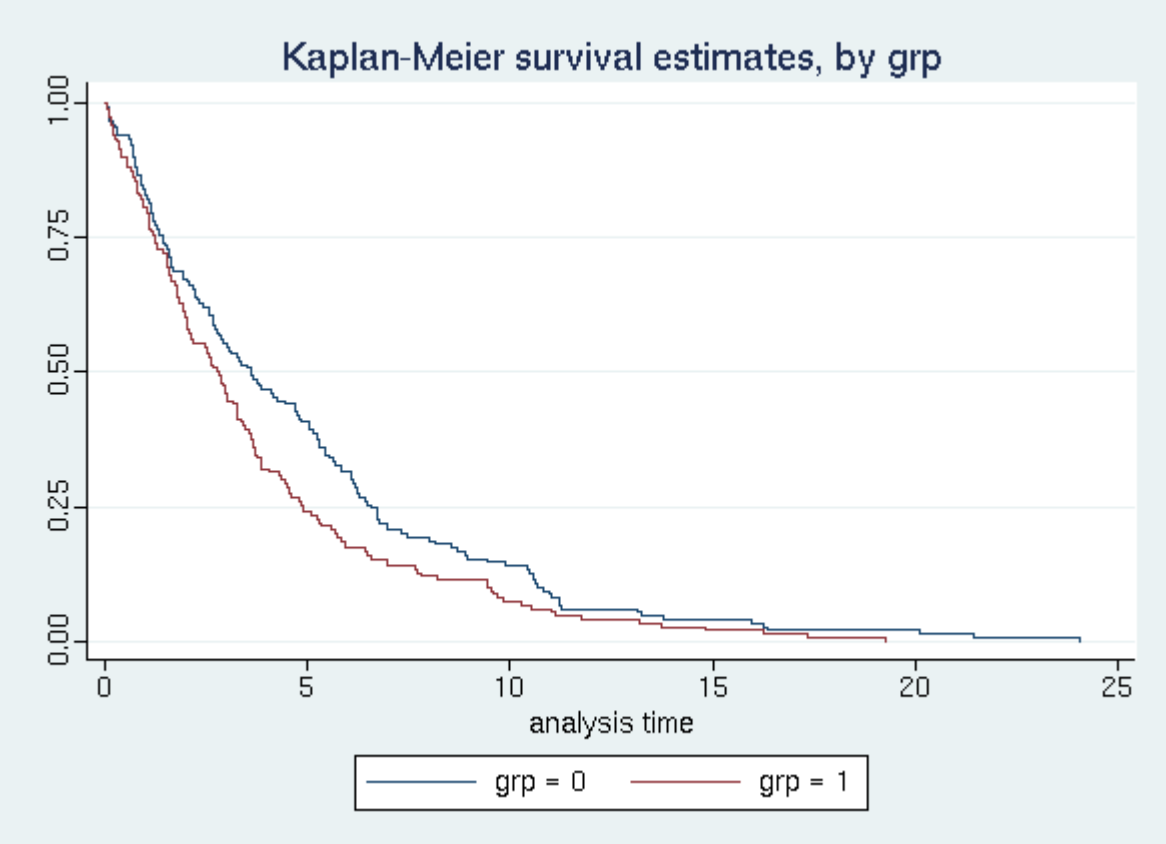
```
regr lt grp

      Source |       SS       df       MS              Number of obs =     300
-------------+------------------------------           F(  1,   298) =    3.61
       Model |  5.06058302     1  5.06058302            Prob > F      =  0.0584
    Residual |  417.740992   298  1.40181541           R-squared     =  0.0120
-------------+------------------------------           Adj R-squared =  0.0087
       Total |  422.801575   299  1.41405209           Root MSE      =   1.184

------------------------------------------------------------------------------
          lt |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         grp |  -.2597584   .1367146    -1.90   0.058    -.5288067    .0092899
       _cons |    1.06905   .0966718    11.06   0.000     .8788044    1.259296
------------------------------------------------------------------------------
predict r,res
graph box r, over(grp)
```

One can determine that a comparison of survivor functions is just $S_1(t)=S_0(\theta t)$ where $\theta=\exp(-\beta_1)$. This result tells us that survival time for those in group 1 is just a scaled version of survival time for those in group 0. The scaling factor is θ . In our example, the estimate of θ is exp(0.2598)= 1.2966. So, for example, survival at year 5 for those in group 1 is estimated to the same as survival at year 5*1.2966 = 6.483 for those in group 0.

These approaches based on log(t) are often called accelerated failure time (AFT) methods.

The direct estimation of the function S(t) in the presence of censoring was developed in the 1950's by Kaplan and Meier:

```
. stset time died
. sts graph, by(grp)
```



Notice that the blue curve [group 0] at year 6.483 is about the same value as the red curve [group 1] at year 5.

Survival functions from Kaplan-Meier methods can be compared as we did with life tables:

```
. sts list,by(grp) at(1 2 3 4 5 6 7 8 9 10 15 20)

         failure _d:  died
   analysis time _t:  t

              Beg.                       Survivor      Std.
    Time      Total      Fail            Function     Error     [95% Conf. Int.]
-------------------------------------------------------------------------------
grp=0
       1        127        24              0.8400    0.0299     0.7709    0.8898
       2        102        25              0.6733    0.0383     0.5920    0.7420
       3         84        18              0.5533    0.0406     0.4702    0.6287
       4         71        13              0.4667    0.0407     0.3852    0.5439
       5         62         9              0.4067    0.0401     0.3278    0.4839
       6         48        14              0.3133    0.0379     0.2409    0.3882
       7         32        16              0.2067    0.0331     0.1461    0.2746
       8         29         3              0.1867    0.0318     0.1290    0.2527
       9         24         5              0.1533    0.0294     0.1011    0.2156
      10         22         2              0.1400    0.0283     0.0902    0.2005
      15          7        15              0.0400    0.0160     0.0165    0.0802
      20          4         3              0.0200    0.0114     0.0055    0.0530
grp=1
       1        122        29              0.8067    0.0322     0.7339    0.8614
       2         91        31              0.6000    0.0400     0.5170    0.6733
       3         70        21              0.4600    0.0407     0.3788    0.5373
       4         49        21              0.3200    0.0381     0.2470    0.3952
       5         37        12              0.2400    0.0349     0.1751    0.3107
       6         27        10              0.1733    0.0309     0.1177    0.2379
       7         22         5              0.1400    0.0283     0.0902    0.2005
       8         19         3              0.1200    0.0265     0.0742    0.1776
       9         18         1              0.1133    0.0259     0.0690    0.1699
      10         12         6              0.0733    0.0213     0.0388    0.1223
      15          4         8              0.0200    0.0114     0.0055    0.0530
      20          1         3                   .         .          .         .
-------------------------------------------------------------------------------
Note:  Survivor function is calculated over full data and evaluated at
       indicated times; it is not calculated from aggregates shown at left.
```

A comparison of survival can be made at any given year using the difference in the estimates and the standard error of the difference. For example, at 5 years, we find

the estimate is : 0.4067 – 0.2400 = 0.1667 the standard error is : $\sqrt{0.0401^2+0.0349^2}=0.0532$

z = 0.1667/0.0532 = 3.1334 p-value= 0.0017

```
disp 2*(1-normal(3.1334)) = .00172794
```

Like with life table comparisons, time specific comparisons from KM methods can be far more illuminating than any omnibus test. There is certain settings when a single test can compare two survivor functions but in order to do so requires critical assumption(s) that need to be checked first.

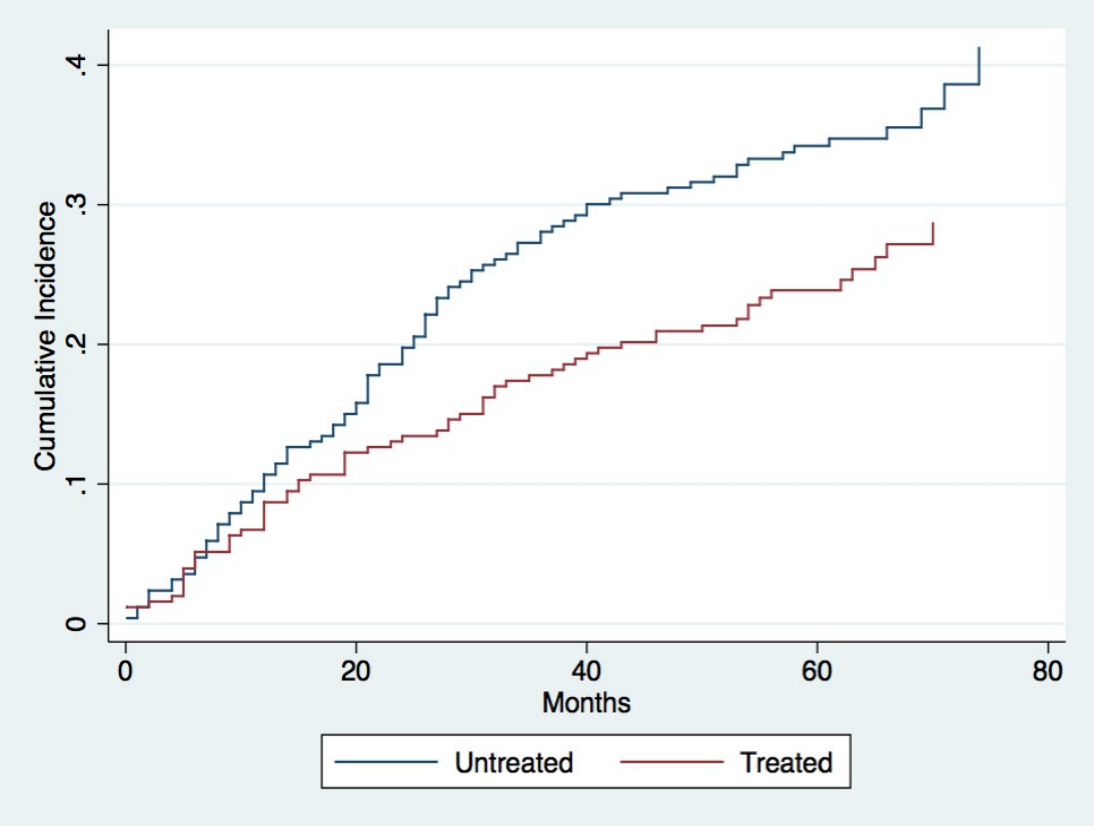
Introduction to Competing Events

With studies involving competing events, the approach is to determine and graph estimates of the Failure function [in the competing events world, also called the cumulative incidence function].

Consider a study [Byar and Green (1980)] of the survival times of 506 patients with prostate cancer who are randomly allocated to a treatment with diethylstilbestrol. 1 The value of the variable status classifies the cause of death as 1 = cancer (the event of interest), 2 = cardiovascular disease, and 3 = other. The patients are considered treated if they received at least 1 mg of diethylstilbestrol daily. In this situation, there are two events competing with the event of interest. The probability of occurrence of each of them has to be separately computed in the two treatment arms. [Use findit stcompet and

download as instructed]

```
. use byar.dta
. stset time, f(status==1)
. stcompet CI = ci hilim = hi lowlim = lo seci = se, compet1(2) compet2(3) by(treatment)
. twoway (line CI time if treat==0 & status==1,connect(stairstep) sort legend(label(1
"Untreated")))(line CI time if treat==1 & status==1,sort connect(stairstep) legend(label(2
"Treated")))
. list time CI se hilim lowlim if treat==0 & status==1 & time ==29
. list time CI se hilim lowlim if treat==1 & status==1 & time ==29
```



Another method [for example, Pintilie 2006 gives the details] now provides us with estimates of the Failure function $F_1(t)$ for the primary event [coded type1] as opposed to the Survivor function $S_1(t)$.

[adapted from the Stata manual] Instead of focusing on the survivor function for the event of interest, P(T > t and event type 1), when competing risks are present you want to focus on the failure function, P(T <= t and event type 1). Part of the rationale for this change in focus is that one will not know what type of event will occur until after it has occurred. It makes more sense to ask "What is the probability of breast cancer within 5 months?" than to ask "What is probability that nothing happens before 5 months, and that when something does happen, it will be breast cancer and not death?"

We can see, in this example, that those treated have an estimated failure function that is not as steep as those untreated [confidence intervals can be added too]

In studies with competing events, it is incorrect to use the Kaplan-Meier method. While 1-KM provides estimates of the Failure function in the presence of censoring, 1-KM is not recommended in the presence of competing events.

https://dspace.ucalgary.ca/bitstream/1880/46804/1/Brar_2008.pdf

Introduction to Hazard

Another crucial notion, that of hazard, is defined next. It turns out to best to prepare the definitions separately for the discrete case and the continuous case.

Hazard: Discrete Time

The hazard function, h(t), is the conditional probability of failure in the ith interval given survival to the previous interval. Using our definitions, we see that $h(t_i)=\frac{p_i}{P_{i-1}}$. Knowledge of any one of p(t), S(t), F(t) or h(t) determines the other three functions. Classic estimation can be accomplished using life table methods.

With more than one group, we can compare time interval specific hazards using the same method illustrated for survival.

Hazard: Continuous Time

As above, the outcome t [= time until event] has a density function: f(t). The area under this curve gives us probability. We continue to consider the distribution function F(t)= P((0,t]) and the survival function S(t)=1-F(t). So F gives us probability up to time t while S gives probability beyond t. These three functions f, F and S are now 'smooth' functions. F and S are not step functions in the continuous world [although we will see below that the estimates of these functions will be steps].

Now we will consider the hazard function $h(t)=\frac{f(t)}{S(t)}$ which is analogous to the discrete case replacing p(t) by f(t). Unfortunately, the hazard is not a conditional probability as in the discrete case. The hazard here can take on any positive value [rather like odds]. It is important here that we speak of the hazard. In the continuous case, the hazard is not probability or risk.

We will also see another definition used widely. It is called the cumulative hazard:

H(t)= - log(S(t))

With a little bit of [ouch] calculus, we can see that h and H are related, indeed $h(t)=\frac{d}{dt}H(t)$.

We will see that estimation of H(t) is quite directly available using methods developed by Nelson and Aalen while estimation of h(t) is more elaborate.

So we now have five functions: f(t), F(t), S(t), h(t) and H(t). Knowledge of any one of these functions determines the other four functions.

Rather like logistic regression where we study log odds, now, in the time-to-event world, we will be studying the log hazard. The 'simplest' log hazard functions are:

Weibull: $h_W(t)=\lambda p t^{p-1}$ $H_W(t)=\lambda t^p$ $\log h_W(t)=\log(\lambda)+\log(p)+(p-1)\log(t)$

where the log of hazard is linear in the log of time. The slope of the line is p-1

Gompertz: $h_G(t)=\lambda e^{\gamma t}$ $H_G(t)=\lambda\gamma^{-1}(e^{\gamma t}-1)$ $\log h_G(t)=\log(\lambda)+\gamma t$

where the log of the hazard is linear in time itself. The slope of the line is γ .

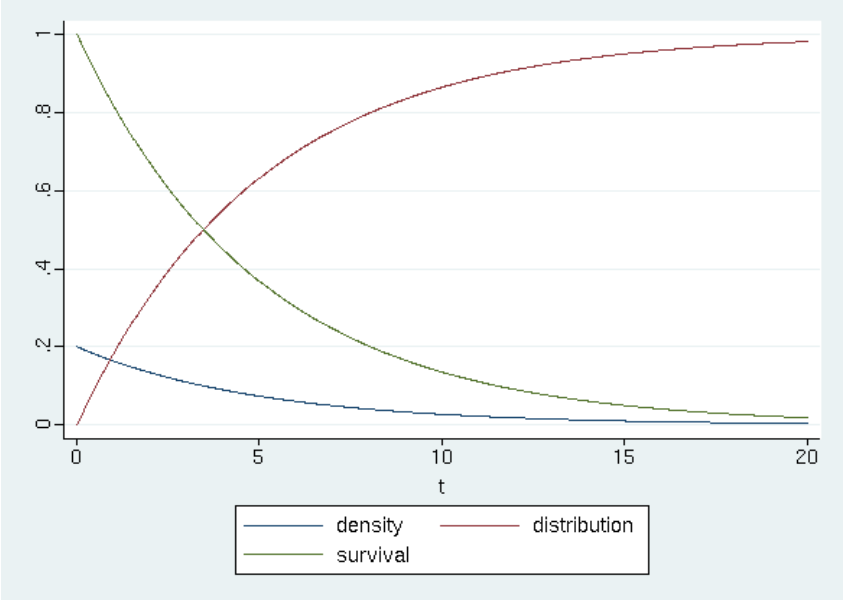
For any given value of p or γ , both of these hazard functions are monotone (constant, always increasing or always decreasing). There are a number of alternative survival functions that provide for

nonmonotone hazard functions (the log-normal hazard and the log-logistic hazard are often cited as is the generalized gamma hazard : all are available for use in Stata or R)

Notice that, for the Weibull, if p=1 we get constant hazard. Similarly, for the Gompertz, if $\gamma=0$ we also get constant hazard. This special case arises with the exponential distribution given by:

$$f(t)=\lambda e^{-\lambda t} \;:\; F(t)=1-e^{-\lambda t} \;:\; S(t)=e^{-\lambda t}:$$
$$h(t)=\lambda \;:\; H(t)=\lambda t \;:\; \log(h(t))=\log\lambda \;:\; \log(H(t))=\log\lambda+\log t$$

Lets try a little example. Suppose we consider a constant hazard of $\lambda=1/5$
Lets consider some probabilities based on this distribution. The probability of surviving for one year is:
$S(1)=e^{-0.2}\approx 0.8187$ while the probability of surviving for 5 years is:



$S(5)=e^{-1}\approx 0.3679$ The median survival time is M where
$S(M)=e^{-0.2M}=0.5$ so that $M=5\log(2)\approx 3.4657$
Now let us suppose that we know that a patient has survived one year. We can ask for the conditional probability that this patient lives for 5 more years (a total of 6 years) given he has lived one year. This is:

$$\frac{S(6)}{S(1)}=\frac{e^{-0.2*6}}{e^{-0.2}}=e^{-0.2*5}=e^{-1}\approx 0.3679$$

So knowledge that a patient has lived one year has no impact on the probability of living another 5 years. This is sometimes called a 'no aging' phenomenon.

Lets now look at this for general λ and an arbitrary condition of say s years.
We can ask for the probability that a patient will lives for u more years (a total of t+u years) given this patient has lived for t years:

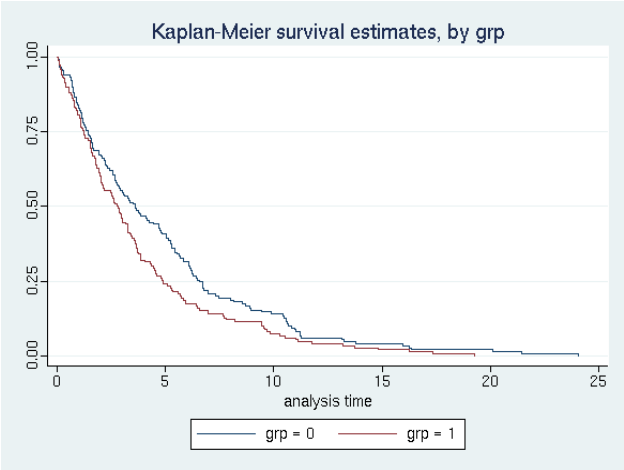
$$\frac{S(t+u)}{S(t)}=\frac{e^{-\lambda(t+u)}}{e^{-\lambda t}}=e^{-\lambda u}=S(u)$$

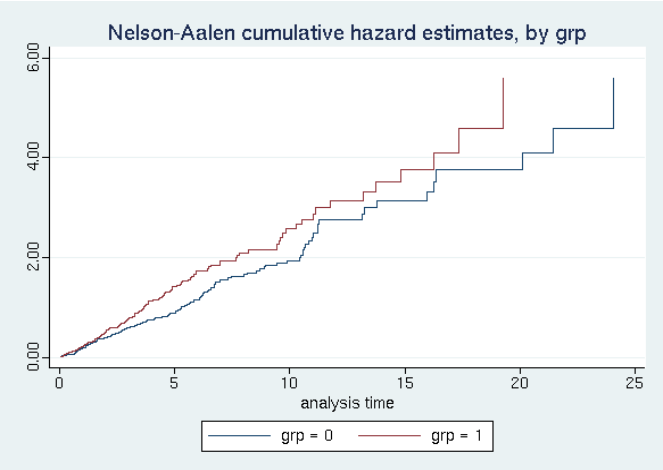
We see that, for the exponential distribution, there is 'no aging'.

This is an implication of the exponential distribution having constant hazard.

An example should help to bring all these concepts into focus. Let us suppose that 2 groups are followed from onset of treatment until death (time t in years: exact date of death known; no censoring or competing events). The estimates of the survival functions and the cumulative hazard functions look like:

```
. use intro_surv_1.dta
. sts graph, by(grp)
```





. sts graph, na by(grp)

From the plot of the estimates of the 2 cumulative hazard curves, it appears reasonable that the hazards are constant (the cumulative hazards are close to lines). In this situation, then, it seems reasonable, for the moment at least, to assume that the 2 hazard curves are constant and hence proportional. If the proportionality assumption is reasonable then we can see that ratio of hazards does not depend on time. This provides support for the comparison of survival curves using the proportional hazard assumption (here, no evidence against the assumption of an 'assumed common' hazard ratio [common across time].

If we are prepared to assume proportional hazards, we can consider the Mantel-Haentzel test here (also called the log-rank test in this context). This test addresses the null hypothesis that the assumed common hazard ratio is one:

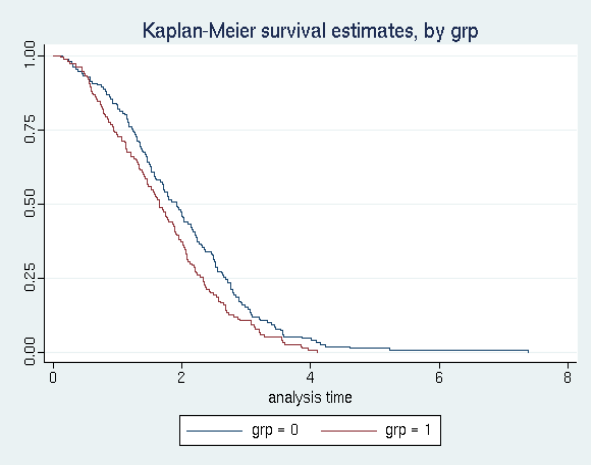
$$H_0: \frac{h_1(t)}{h_0(t)} = 1 \text{ or } H_0: \log h_1(t) - \log h_0(t) = 0$$

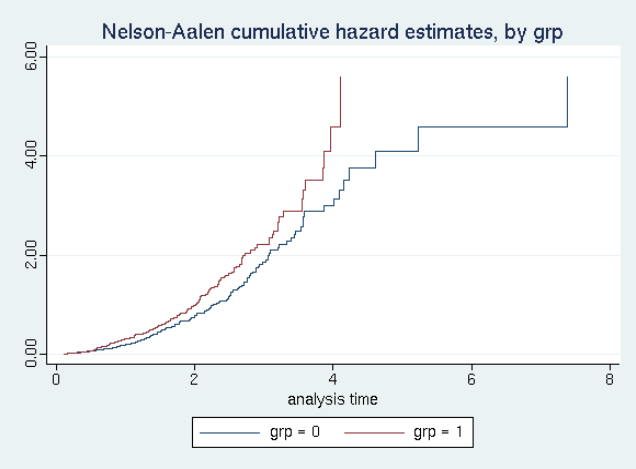
Or that the assumed common difference in log hazards is zero.

```
. sts test grp

Log-rank test for equality of survivor functions

      |   Events         Events
grp   |  observed       expected
------+-------------------------
0     |       150          169.74
1     |       150          130.26
------+-------------------------
Total |       300          300.00

           chi2(1) =       5.39
           Pr>chi2 =     0.0203
```

A second example:

```
. use intro_surv_2.dta
. sts graph, by(grp)
. sts graph, na by(grp)
```





Here the cumulative hazard functions are clearly not linear, telling us that the hazard is not constant. Indeed we are seeing, cumulative hazards that look to be quadratic so that, here, we may have a linear hazard function for both groups and then, again, the assumption of proportional hazards might be reasonable. More on this matter is coming up. Anyhow, if we are prepared to assume that the two hazard functions are proportional, we can consider the log-rank test with the same null hypothesis as before.

```
. sts test grp
Log-rank test for equality of survivor functions

      |   Events         Events
grp   |  observed       expected
------+-------------------------
0     |       150         170.81
1     |       150         129.19
------+-------------------------
Total |       300         300.00

           chi2(1) =       6.04
           Pr>chi2 =     0.0140
```

Subhazard in the World of Competing Events

We defined the event specific Failure function $F_1(t)$ and now, analogous to the hazard definitions above, we have the cumulative subhazard function:

$$\bar{H}_1(t) = -\log(1 - F_1(t))$$

and the subhazard function:

$\bar{h}_1(t) = \frac{d\,\bar{H}_1(t)}{dt}$ or $\bar{h}_1(t) = \frac{f_1(t)}{1 - F_1(t)}$. f_1 and F_1 are the density and failure of [what is called] the subdistribution. The area under f_1 is not one and F_1 does not reach one on the right.

For now, it may be best to note the relationship between the subhazard and the Failure function:

$$F_1(t) = 1 - \exp(-\bar{H}_1(t))$$

Using this relationship, we will see later that a study the log of the subhazard, then yields estimates of the Failure function.