Hazard function

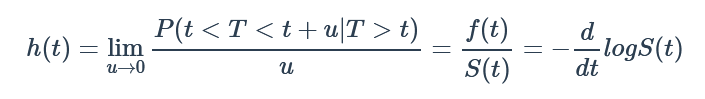
Survival function

A picture containing diagram

Description automatically generated

Survival function is non-increasing with t and equals 1 at t = 0.

Another important term in survival analysis is **hazard h(t)** – also known as the force of mortality, i.e., the probability that an individual who is still observed at time t has an event in small interval after t:



The numerator in the first part of the above expression is the conditional probability that the event will occur in the small interval after *t* (*[t, t+u)*) given that it has not occurred before, and the denominator is the width of the interval (*u*). The second part of the formula says that the hazard rate of mortality at time *t* equals the density of deaths at *t*, divided by the probability of membership lasting to that point in time without dying.

The area under the curve for hazard *h(t)* is a **cumulative hazard function (*H(t)*)**:

Logo

Description automatically generated

Text

Description automatically generated

Therefore, knowing any one of the functions *S(t)*, *H(t)*, or *h(t)* allows one to derive the other two functions. These three functions are different ways of describing the same distribution for the time to event.

Proportional Hazards

Diagram

Description automatically generated

where *h*0(*t*) is the baseline hazard function (when all predictors equal zero), *X* - matrix of features (not including a constant!), *β* - regression coefficients (weights). Here *exp*(*x*‘*iβ*) is the relative risk associated with the set of characteristics x\_i for the i-th individual in comparison with the risk of an individual with all characteristics equal zero.

While AFT (Accelerated Failure Models) models assumes that predictors have a multiplicative effect on the time to event and additive effect for the log time to event, proportional hazards models assumes the same for the hazard and log hazard.

*h*(*t*|*X*) = *h*0(*t*)*exp*(*Xβ*) ⇒ *log*(*h*(*t*|*X*)) = *log*(*h*0(*t*))+*Xβ*

It is important to emphasize that the relative risk does not depend on time, i.e. it is constant in time for the same pair of values of any feature, so hazards are proportional independent of time.

We can use some specific distribution for the baseline survival or hazard function – and then we get a **parametric PH model**, or do not make such assumptions about the baseline functions – and then get a **semi-parametric PH model**, often called a Cox PH model after its author.

**Semi-parametric (Cox) PH models** are more widely used than parametric PH models, because they do not have parametric assumptions concerning baseline functions - just the effect of covariates on hazard. Cox argued that when the PH assumption holds, information about baseline hazard function is not very useful in estimating the parameters of primary interest (*β*). By special conditioning in formulating the log likelihood function, Cox showed how to derive a valid estimate of *β* that does not require estimation of baseline hazard as it is dropped out of the new likelihood function. Cox’s derivation focuses on using the information in the data that relates to the relative hazard function exp(*Xβ*) ([Harrell 2015](https://rstudio-pubs-static.s3.amazonaws.com/850319_ac024ddba2a8430192318fe830778e3f.html#ref-harrell_regression_2015)).

As a result, estimation of the Cox PH model is based on the maximization of the partial likelihood:

Logo, company name

Description automatically generated

or, equivalently, minimization of the negative log partial likelihood −*logL*

. Newton-Rapshon technique for iterative estimation is often used here.

After estimating the Cox PH we can get a linear prediction for each individual as *x*‘*iβ*̂

and use it to rank them by relative risk of event.