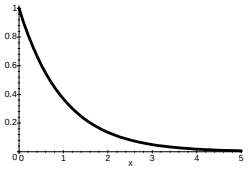
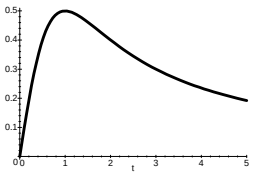
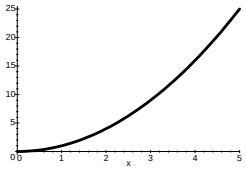
Shapes of hazard rates



Mortality of patients with myocardial infarction.



Divorce rates.
Mortality of cancer patients.
Incidence of childhood leukemia.



General mortality.
Incidence of most cancers.

How can the different patterns be interpreted?

We need to distinguish between the individual hazards and the population hazard

Heterogeneity in survival analysis

- Differences between individuals are of two kinds:
  - those observed through covariates
  - those that are unobserved
- The latter ones are usually disregarded in survival analysis
- This may lead to distortions as explained in the *frailty* theory
- There is a large literature on frailty theory, but we will review just a few basic issues

Population hazard and individual hazard

- We will distinguish between the population hazard rate and the individual hazard rates
- The population hazard rate is influenced by selection: those with highest risk experience the event early
- The shape of the population hazard may be entirely different from that of the individual hazards
- Hence the population hazard can not be interpreted as giving information on individual development in risk

The proportional frailty model

We assume that the heterogeneity between individuals may be described by a frailty variable Z

The frailty variable is a non-negative random variable, with large values of Z corresponding to "frail" individuals

It is common to assume that the frailty has a multiplicative effect on the hazard, i.e.

$$\alpha(t \mid Z) = Z \cdot \alpha(t)$$

Here $\alpha(t)$ is the individual baseline hazard (corresponding to Z=1) and $\alpha(t \mid Z)$ is the individual hazard for an individual with frailty Z

Note that the frailty is not observed

The most common choice of frailty distribution is the gamma distribution with density ($z > 0$)

$$f(z) = \frac{\nu^{\eta}}{\Gamma(\eta)} z^{\eta-1} \exp(-\nu z)$$

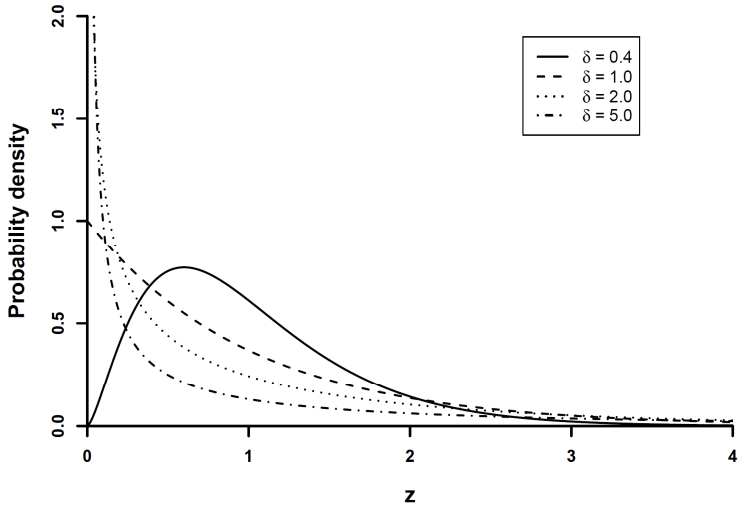
It is well known that the gamma distribution has mean η/ν and variance η/ν^2

We will often assume that the frailties have mean equal to 1

For the gamma distribution this implies that $\eta = \nu$

Then the variance becomes $\delta = \nu^{-1}$

Gamma distributions with mean 1 for some values of the variance δ



Laplace transform

The Laplace transform is a convenient tool to study the multiplicative frailty model

For a positive random variable Z the Laplace transform is given by

$$\mathscr{L}(c) = \mathrm{E}(e^{-cZ})$$

The Laplace transform is closely related to the moment generating function

$$\mathscr{M}(s) = \mathrm{E}(e^{sZ})$$

It is well known that for the gamma distribution with mean η/ν and variance η/ν^2, the moment generating function takes the form

$$\mathscr{M}(s) = \left(\frac{1}{1 - s/\nu}\right)^{\eta}$$

Thus the Laplace transform becomes

$$\mathscr{L}(c) = \mathscr{M}(-c) = \left(\frac{1}{1 + c/\nu}\right)^{\eta}$$

In particular for the gamma distribution with mean 1 (i.e. $\eta = \nu$) and variance $\delta = \nu^{-1}$ the Laplace transform takes the form

$$\mathscr{L}(c) = \left(\frac{1}{1 + c/\nu}\right)^{\nu} = (1 + c/\nu)^{-\nu} = (1 + \delta c)^{-1/\delta}$$

Population survival function

Consider a population where the heterogeneity is described by the proportional frailty model

Let T be the survival time of a randomly selected individual from the population

Then $S(t \mid Z) = P(T > t \mid Z) = \exp\{-Z\,A(t)\}$

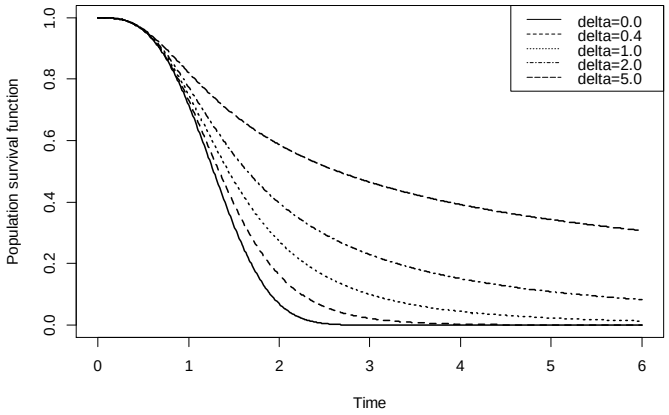
where $A(t) = \int_0^t \alpha(u)du$

The population survival function is given by

$$S(t) = P(T > t) = \mathrm{E}\{I(T > t)\}$$
$$= \mathrm{E}\big(\mathrm{E}\{I(T > t) \mid Z\}\big) = \mathrm{E}\{P(T > t \mid Z)\}$$
$$= \mathrm{E}(e^{-ZA(t)}) = \mathscr{L}(A(t))$$

If frailty is gamma distribution with mean 1 and variance δ the population survival function becomes

$$S(t) = \mathscr{L}(A(t)) = \{1 + \delta\,A(t)\}^{-1/\delta}$$



Population hazard

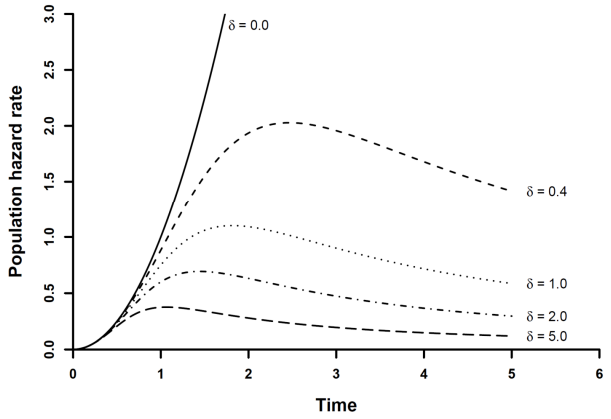
The population hazard becomes

$$\mu(t) = \frac{-S'(t)}{S(t)} = \frac{-\frac{d}{dt}\mathscr{L}(A(t))}{\mathscr{L}(A(t))} = \alpha(t)\frac{-\mathscr{L}'(A(t))}{\mathscr{L}(A(t))}$$

If frailty is gamma distribution with mean 1 and variance δ the population hazard rate is given by

$$\mu(t) = \alpha(t)\frac{-\mathscr{L}'(A(t))}{\mathscr{L}(A(t))} = \alpha(t)\frac{(1+\delta\,A(t))^{-\frac{1}{\delta}-1}}{(1+\delta\,A(t))^{-\frac{1}{\delta}}}$$

$$= \frac{\alpha(t)}{1+\delta A(t)}$$



One sees that the population hazard is "pulled down" with a strength determined by δ

Estimating frailty

- For survival data where only a single event is available for each individual, the frailty effect is not identifiable unless we assume a specific form of the individual baseline hazard rate $\alpha(t)$
- Frailty models for survival data may be speculative, but the are useful for understanding why the population hazard may have different shapes
- Estimation of frailty is more relevant for clustered survival data and recurrent event data (repeated events)

Examples of clustered survival data

Examples 1.9 and 7.2: Duration of amalgam fillings in teeth for 32 patients

Patient number	Age at inclusion (years)	Gender (male/female)	Number of fillings inserted*	Number of fillings at 10 years†
1	61	F	10	6
2	38	F	37	14
3	62	F	21	13
4	62	F	18	14
5	61	F	26	11
6	75	F	12	4
7	59	F	19	13
8	62	F	17	14
9	56	F	12	5
10	49	F	26	12
11	37	F	25	10
12	49	F	9	2
13	31	F	9	2
14	77	F	8	2
15	43	F	14	11
16	34	F	17	12
17	50	M	27	13
18	64	M	12	1
19	50	M	7	5
20	50	M	14	5
21	65	M	4	1
22	44	M	18	9
23	64	M	32	11
24	58	M	7	5
25	45	M	23	8
26	34	M	38	13
27	62	M	18	6
28	34	M	7	2
29	45	M	30	18
30	71	M	10	3
31	72	M	16	5
32	43	M	23	9

* Number of fillings inserted during period of observation
† Number of fillings under observation for at least 10 years

Example: Litter-matched rats

The data consist of 50 litters of female rats with 3 rats in each litter

One rat in each litter received a potentially tumorgenic treatment, the other two were controls

The time (in weeks) until tumor occurrence was observed for each rat

Cencoring was due to death without tumor or end of study (at 104 weeks when the rats still alive were sacrificed)

Data for 6 of the 50 litters

(T is occurrence of tumor, D is death or sacrifice, i.e. censoring)

Litter no.	Drug-treated	Control 1	Control 2
01	101-D[a]	49-T	104-D
03	104-D	102-D	104-D
05	104-D	104-D	104-D
07	77-D	97-D	79-D
09	89-D	104-D	104-D
11	88-T	96-T	104-D

A shared frailty model for clustered survival data

Assume that we have data from m independent clusters

In the i-th cluster there are n_i units

The uncensored survival times of the units in the i-th cluster are denoted T_{ij} $(j = 1,...,n_i)$

$T_{i1},....,T_{in_i}$ are *dependent* survival times

To handle the dependence, we will assume that the units in the i-th cluster share the same frailty Z_i

We will assume that *given* Z_i, the survival times $T_{i1},....,T_{in_i}$ are independent and have hazard rates $Z_i \cdot \alpha_{ij}(t)$, where $\alpha_{ij}(t)$ may depend on covariates (for the clusters and/or the units)

In the i-th cluster we observe censored survival times $\tilde{T}_{ij}$ and censoring indicators D_{ij} $(j = 1,...,n_i)$

The frailty variables $Z_1,....,Z_m$ for the *m* clusters are assumed to be independent and identically distributed (e.g. gamma distributed)

It is common to assume $\mathrm{E}(Z_i) = 1$

We will derive the marginal likelihood for the data (i.e. the likelihood based on the distribution of the observable data)

Data for the i-th cluster (in addition there may be covariates)

$$H_i = \left\{\tilde{T}_{ij}, D_i ; j = 1,....,n_i\right\}$$

Conditional in frailty, the likelihood contribution for the i-th cluster becomes (cf. chapter 5)

$$P(H_i \mid Z_i) = \prod_{j=1}^{n_i}\left\{\left(Z_i\, \alpha_{ij}(\tilde{T}_{ij})\right)^{D_{ij}} \exp\left(-Z_i\, A_{ij}(\tilde{T}_{ij})\right)\right\}$$

where $A_{ij}(t) = \int_0^t \alpha_{ij}(u)du$

Note that

$$P(H_i \mid Z_i) = \prod_{j=1}^{n_i}\left\{\left(\alpha_{ij}(\tilde{T}_{ij})\right)^{D_{ij}}\right\} \cdot Z_i^{D_{i\bullet}} \exp\left(-Z_i V_i\right)$$

where $D_{i\bullet} = \sum_{j=1}^{n_i} D_{ij}$ and $V_i = \sum_{j=1}^{n_i} A_{ij}(\tilde{T}_{ij})$

We obtain the contribution to the likelihood for the i-th cluster by integrating over the frailty distribution:

$$P(H_i) = \prod_{j=1}^{n_i}\left\{\left(\alpha_{ij}(\tilde{T}_{ij})\right)^{D_{ij}}\right\} \cdot E_{Z_i}\left\{Z_i^{D_{i\bullet}} \exp\left(-Z_i V_i\right)\right\}$$

Remember the Laplace transform: $\mathscr{L}(c) = E(e^{-cZ})$

Note that the *r* -th derivative becomes:

$$\mathscr{L}^{(r)}(c) = (-1)^r \mathrm{E}(Z^r e^{-cZ})$$

Hence we may write

$$P(H_i) = \prod_{j=1}^{n_i}\left\{\left(\alpha_{ij}(\tilde{T}_{ij})\right)^{D_{ij}}\right\} \cdot (-1)^{D_{i\bullet}} \mathscr{L}^{(D_{i\bullet})}(V_i)$$

The total likelihood is given by

$$L = \prod_{i=1}^{m} P(H_i)$$

so the log-likelihood becomes

$$\log L = \sum_{i=1}^{m}\left[\sum_{j=1}^{n_i} D_{ij} \log\left(\alpha_{ij}(\tilde{T}_{ij})\right) + \log\left\{(-1)^{D_{i\bullet}} \mathscr{L}^{(D_{i\bullet})}(V_i)\right\}\right]$$

If $\alpha_{ij}(t)$ is given a parametric specification, the log-likelihood may be maximized with respect to the parameters of $\alpha_{ij}(t)$ (which may include parameters for the baseline hazard and regression coefficients) and the parameters of the frailty distribution

Standard ML-results apply

Example 7.2: Duration of amalgam fillings

Each person is a cluster, and the fillings are the units

We fit a gamma frailty model (where the frailties have mean 1 and variance δ) with no covariates, assuming a Weibull baseline hazard:

$$\alpha(t \mid Z_i) = Z_i \cdot \alpha(t) = Z_i \cdot bt^{k-1}$$

ML-estimates (with standard errors):

$$\log(\hat{b}) = -4.21\ (0.25)$$
$$\hat{k} = 0.43\ (0.10)$$
$$\hat{\delta} = 0.85\ (0.31)$$

The R library "parfm" may be used to fit frailty models with parametric baseline hazard

Default is a Weibull baseline parameterized as $\alpha(t) = \lambda\rho t^{\rho-1}$

When using a gamma frailty, it is assumed to have mean 1 and variance θ (corresponding to δ above)

See the documentation for other baseline hazards and frailty distributions

A detailed description of the package is given in a paper at http://www.jstatsoft.org/v51/i11/

Example: Litter-matched rats

Each litter is a cluster, and the rats are the units

We fit a gamma frailty model (where the frailties have mean 1 and variance θ) with treatment as covariate, assuming a Weibull baseline hazard:

$$\alpha_{ij}(t \mid Z_i) = Z_i \cdot \alpha_{ij}(t) = Z_i \cdot \lambda\rho t^{\rho-1} \exp(\beta x_{ij})$$

Here $x_{ij} = 1$ if rat j in litter i is treated, $x_{ij} = 0$ otherwise

ML-estimates (with standard errors):

$$\hat{\rho} = 3.93\ (0.57) \qquad \hat{\lambda} = 0.020\ (0.009)$$
$$\hat{\beta} = 0.91\ (0.32) \qquad \hat{\theta} = 0.49\ (0.47)$$

Testing $H_0 : \delta = 0$ vs $H_A : \delta > 0$

We will use the likelihood ratio test, to test if the there is an effect of frailty

This corresponds to testing if the frailty variance δ (or θ) is 0

The usual properties for the likelihood ratio test do not apply in this situation, since the null hypothesis is at the boundary of the parameter space

One may show that in such situations two times the difference in log-likelihoods is approximately distributed as

$$\tfrac{1}{2}\chi_0^2 + \tfrac{1}{2}\chi_1^2$$

To get the correct P-value, we should therefore simply halve the P-value we obtain from the usual likelihood ratio test

Example: Litter-matched rats

The log-likelihood for the model with frailty is -83.423

The model without frailty has log-likelihood -84.277

The likelihood ratio statistic takes the value
2*(-83.423 +84.227) =1.608

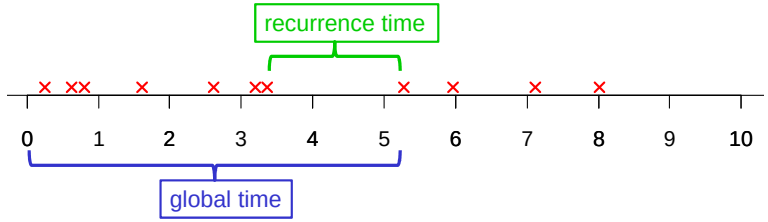
The P-value becomes
0.5*0.2048=0.102

Thus there is not a significant litter effect for the rat data

Recurrent event data

For each of *m* individuals we are observing repeated occurrences of an event (e.g epileptic seizures, heart attacks)

Data for one individual (events marked with x):



For modelling one may use global time (time since start) or recurrence time (time since last event)

When using recurrence time to model recurrent event data, a much used model is a renewal process where the times between events for an individual are assumed iid

Then the data may be treated as clustered survival data with one cluster for each individual

Example of clustered data from recurrent events

Examples 1.11 and 7.1: Movements of the small bowl

Individual	Completely observed periods	Censored period
1	112 145 39 52 21 34 33 51	54
2	206 147	30
3	284 59 186	4
4	94 98 84	87
5	67	131
6	124 34 87 75 43 38 58 142 75	23
7	116 71 83 68 125	111
8	111 59 47 95	110
9	98 161 154 55	44
10	166 56	122
11	63 90 63 103 51	85
12	47 86 68 144	72
13	120 106 176	6
14	112 25 57 166	85
15	132 267 89	86
16	120 47 165 64 113	12
17	162 141 107 69	39
18	106 56 158 41 41 168	13
19	147 134 78 66 100	4

We fit a gamma frailty model for the recurrence times (where the frailties have mean 1 and variance δ) with no covariates, assuming a Weibull baseline hazard:

$$\alpha(t \mid Z_i) = Z_i \cdot \alpha(t) = Z_i \cdot bt^{k-1}$$

ML-estimates (with standard errors):

$$\log(\hat{b}) = -10.0 \ (1.0)$$

$$\hat{k} = 2.28 \ (0.22)$$

$$\hat{\delta} = 0.15 \ (0.12)$$

To test $H_0 : \delta = 0$ vs $H_A : \delta > 0$ we may use the one-sided likelihood ratio test

Twice the difference in log-likelihoods becomes 2.58 corresponding to a P-value of 5.4 %