Frailty models

STK4080 H16

1. Concept of frailty (Ch. 6 & 6.1)
2. Marginal hazard in frailty models (Ch. 6.2.1 & 6.2.2)
3. Recurrent and cluster data (Ch. 7 & 7.1)
4. Shared frailty models (Ch. 7.2)
5. Likelihood derivations (Ch.7.2.2)

Frailty

A frailty is a latent (unobserved) random variable Z used in survival analysis assuming that the hazard given Z equals

$$\alpha(t|Z) = Z\alpha(t)$$

for some basic rate $\alpha(t)$. The frailty Z captures heterogeneity in the population.

Frailties can alternatively be described as

- mixing distributions
- random components

They can describe situations where what is observed on a populations level may differ from what goes on on the individual level.

Modelling frailties

Typically some parametric model is assumed for the Z and a common choice is the gamma distribution with density

$$Z \sim g(z|\eta,\nu) = \frac{\nu^\eta}{\Gamma(\eta)} z^{\eta-1} \exp(-\nu z); \qquad z > 0$$

with expectation $\mathrm{E}[Z] = \frac{\eta}{\nu}$ and variance $\mathrm{Var}[Z] = \frac{\eta}{\nu^2}$.

To avoid over-parametrization of the model $Z\alpha(t)$ one typically restricts $\mathrm{E}[Z] = 1$, thus $\nu = \eta$ and $\mathrm{Var}[Z] = \frac{1}{\nu} = \delta$.

Note that with $\mathrm{Var}[Z] = \delta \to 0$, the model converges to $T \sim \alpha(t)$, thus a model without frailty.

The degree of heterogeneity is then described by δ.

The marginal survival function

Conditional on Z we have the survival function

$$S(t|Z) = \mathrm{P}(T > t|Z) = \exp(-ZA(t))$$

where $A(t) = \int_0^t \alpha(s)ds$ is the cumulative hazard.

It follows that the marginal or population survival function becomes

$$S(t) = \mathrm{P}(T > t) = \int_0^\infty \exp(-zA(t))g(z)dz$$

where $g(z)$ is the density of the frailty Z.

After obtaining $S(t) = \exp(-M(t))$ where $M(t) = \int_0^t \mu(s)ds$ we obtain the marginal (population) hazard $\mu(t)$ by

$$\mu(t) = \frac{d}{dt}[-\log(S(t))]$$

Laplace transform

The Laplace transform of a random variable Z with density $g(z)$ is given by

$$\mathcal{L}_Z(c) = \mathrm{E}[\exp(-cZ)] = \int \exp(-cz)g(z)dz$$

whenever this integral exists. Note that $\mathcal{L}_Z(c) = M_Z(-c)$ where $M_Z(t) = \mathrm{E}[\exp(Zt)]$ is the moment generating function of Z.

It is convenient to use the Laplace transform in the context of frailty models because the marginal survival function becomes

$$S(t) = \mathrm{P}(T > t) = \int \exp(-zA(t))g(z)dz = \mathcal{L}_Z(A(t))$$

which leads to a marginal hazard

$$\mu(t) = \frac{d}{dt}[-\log(S(t))] = \alpha(t)\frac{-\mathcal{L}'_Z(A(t))}{\mathcal{L}_Z(A(t))}$$

Marginal survival and hazard with gamma-frailty

When $Z \sim gamma(\nu, \nu)$ and $\delta = 1/\nu$ we get a Laplace-transform

$$\mathcal{L}_Z(c) = \frac{1}{(1+\delta c)^{1/\delta}}.$$

Thus the marginal survival function becomes

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

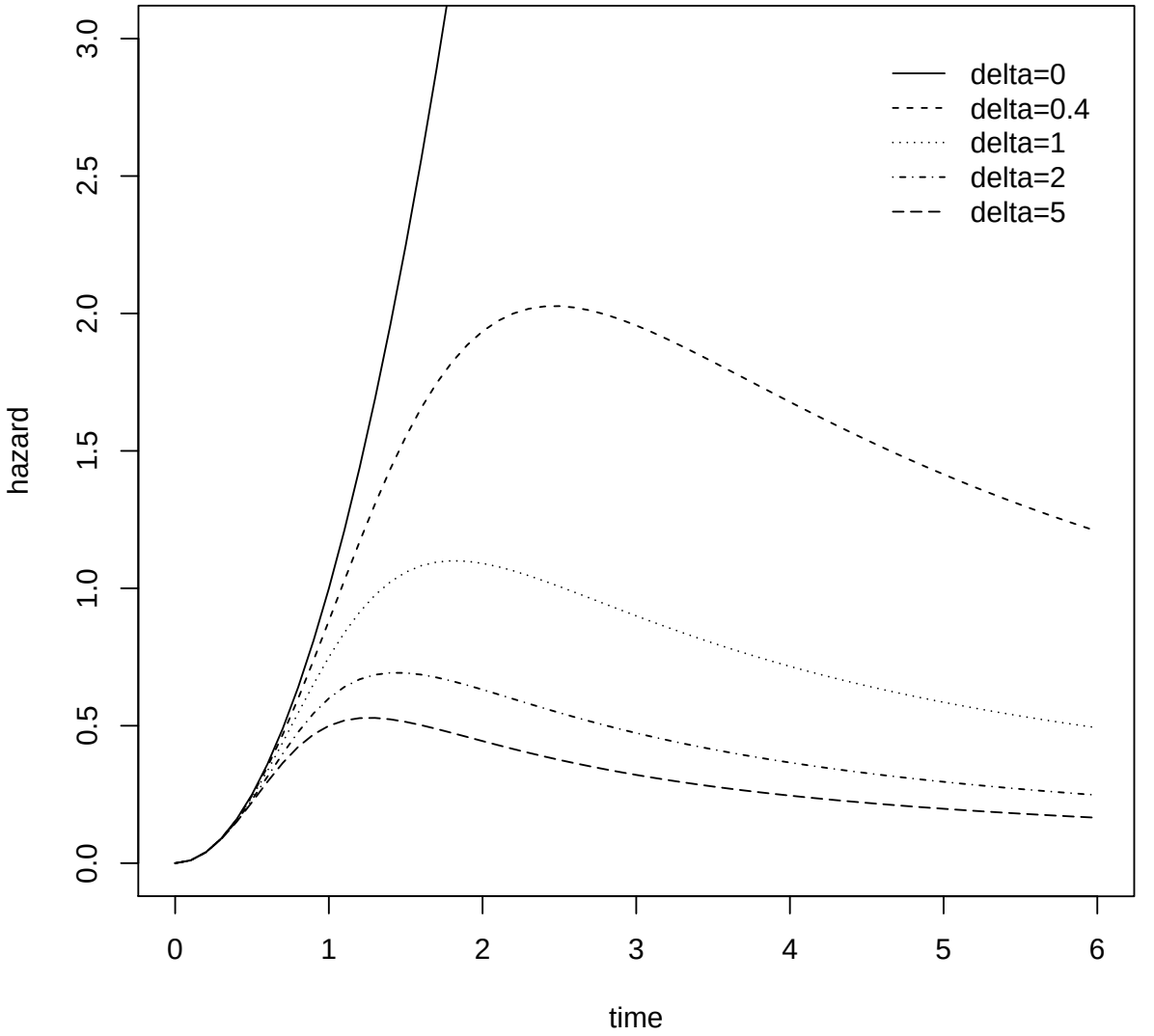
whereas we get the marginal hazard by

$$\mu(t) = \frac{d}{dt}[-\log(S(t))] = \frac{\alpha(t)}{1+\delta A(t)}$$

As shown on the next page the shape of the marginal hazard can look quite different from the basic hazard $\alpha(t)$.

Example: $\alpha(t) = t^2$, Weibull, increasing hazard

With $\alpha(t) = t^2$ we get $A(t) = t^3/3$ and $\mu(t) = \frac{\alpha(t)}{1+\delta A(t)} = \frac{t^2}{1+\delta t^3/3}$



Summarizing

- Even though individual hazards was increasing, the population hazard had a maximum
- Can be explained by the "frail", i.e. high Z, fail first
- After a while the remaining population is relatively robust

This may be an explanation for phenomena with unimodal hazard

- Divorce rates typically have a maximum around 7 years after marriage ("7 year itch")
- Incidence of testicular cancer declines after around 35 years
- Schizophrenia sets in among teenagers/young adults

However, there may be other mechanisms leading to hazards with a maximum.

Multivariat survival

We can say more about the possibility of frailties in the presence of multivariate survival data.

We will discuss two types of multivariate survival data

- Recurrent (serial) times for one individual
- Clustered survival data

Examples of the first may be the times

- between individual ear infections
- between individual admissions to psychiatric wards

Examples of the second could be times to

- asthma diagnosis for two twins (or more siblings)
- failure of amalgam fillings for one individual

Shared frailty model

Let T_{ij} be the times until event j for

- for individual i (recurrent data)
- cluster i

We assume a shared frailty Z_i for each event $j = 1, \ldots, n_i$ where n_i are the number of recurrences / individuals in cluster i.

Conditional on Z_i we assume hazards $Z_i\alpha(t)$ for $T_{i1}, \ldots, T_{in_i}$

Furthermore we assume that independent censoring between T_{ij} and that clusters of data are independent.

With censoring times C_{ij} let $\tilde{T}_{ij} = \min(T_{ij}, C_{ij})$ be the right-censored survival times and $D_{ij} = I(\tilde{T}_{ij} = T_{ij})$ be indicators of events. This are combined into $H_i = (\tilde{T}_{i1}, D_{i1}, \ldots, \tilde{T}_{in_i}, D_{in_i})$.

Likelihood for multivariate frailty data

Conditionally on the frailty Z_i we get the likelihood contribution from individual / cluster i as

$$\begin{aligned} P(H_i|Z_i) &= \prod_{j=1}^{n_i} (Z_i \alpha(\tilde{T}_{ij}))^{D_{ij}} \exp(-Z_i A(\tilde{T}_{ij})) \\ &= (\prod_{j=1}^{n_i} \alpha(\tilde{T}_{ij})^{D_{ij}}) Z_i^{D_{i\bullet}} \exp(-Z_i V_i) \end{aligned}$$

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

The unconditional likelihood contributions can thus be written

$$P(H_i) = \mathrm{E}_{Z_i} P(H_i|Z_i) = (\prod_{j=1}^{n_i} \alpha(\tilde{T}_{ij})^{D_{ij}}) \mathrm{E}_{Z_i} \left[Z_i^{D_{i\bullet}} \exp(-Z_i V_i) \right]$$

and the total log-likelihood as

$$logL = \sum_{i=1}^{n} \log(P(H_i))$$

A property of the Laplace-transform is that its r-th derivative can be written

$$\mathcal{L}_Z^{(r)}(c) = \mathrm{E}_Z\left[(-Z)^r \exp(-cZ)\right].$$

The log-likelihood can thus be expressed through these as

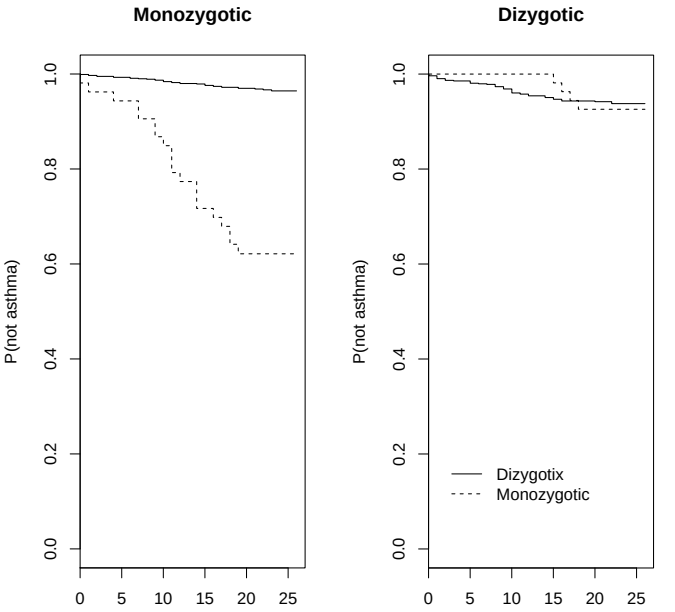
$$logL = \sum_{i=1}^{n}\left[\sum_{j=1}^{n_i} D_{ij}\log(\alpha(\tilde{T}_{ij})) + \log((-1)^{D_{i\bullet}}\mathcal{L}_{Z_i}^{(D_{i\bullet})}(V_i))\right]$$

In particular if $Z_i \sim gamma(1/\delta, 1/\delta)$, i.e. $\mathrm{E}Z_i = 1$ and $\mathrm{Var}Z_i = \delta$ we saw that $\mathcal{L}_Z(c) = [(1+\delta c)]^{-1/\delta}$ which can be plugged in to $logL$ after deriving the r-th derivative of $\mathcal{L}_Z(c)$ (see ABG, pg.279).

Example: Asthma Norwegian twins (Harris et al., 1997)

Age of diagnosis of asthma before age at investigation age 16-25 years was recorded for 527 monozygotic girl twin pairs and 442 dizygotic girl twin pairs.

The plots show Kaplan-Meier estimators of not being diagnosed with asthma given that the co-twin had asthma among monozygotic and dizygotic twins.



Clearly larger risk if a monozygotic co-twin had asthma.

Example: Asthma Norwegian twins, contd.

This would more appropriately be analyzed using shared frailty models.

Will use function `parfm` from R-library `parfm`. Here the Weibull distribution is modelled as $\alpha(t;\rho,\lambda) = \lambda t^{\rho-1}$ and the gamma-variance is denoted $\theta = \delta$.

The output on next slide shows that the standard error of $\hat{\theta} = \hat{\delta}$ is much smaller than the actual estimate for monozygotic twins, whereas it is larger for dizygotic twin.

This is consistent with a genetic effect (but could also be influenced by larger correspondence in environment for monozygotes).

Example: Asthma Norwegian twins, output

```
# Monozygotic girl twins
> parfm(Surv(tid,ast)~1,cluster="par",frailty="gamma",data=j1)

Frailty distribution: gamma
Baseline hazard distribution: Weibull

       ESTIMATE    SE
theta     9.239 3.392
rho       1.022 0.138
lambda    0.003 0.001

# Dizygotic girl twins
> parfm(Surv(tid,ast)~1,cluster="par",frailty="gamma",data=j2)

Frailty distribution: gamma
Baseline hazard distribution: Weibull

       ESTIMATE    SE
theta     0.149 0.734
rho       0.710 0.096
lambda    0.007 0.002
```

Asthma Norwegian twins: In the actual paper

we did not use frailty analysis, but clustered Cox-regression (Chapter 8, not curriculum). Status of the co-twin was covariate.

Such data are dependent, but the dependency can be accounted for by a robust (sandwich estimator) variance.

```
# Monozygotic girl twins
> coxph(Surv(tid, ast) ~ asttv + cluster(par), data = j2)

       coef exp(coef) se(coef) robust se    z   p
asttv 0.183     1.201    0.520     0.715 0.26 0.8

n= 884, number of events= 54

# Dizygotic girl twins
> coxph(Surv(tid, ast) ~ asttv + cluster(par), data = j1)

        coef exp(coef) se(coef) robust se    z       p
asttv  2.638    13.979    0.284     0.396 6.65 2.9e-11

n= 1054, number of events= 53
```