Interview

for

Statistical Analyst Position GGHS

Maciej J. Dańko

MPIDR January 2017

Selected major research projects of 2016 (chronologically)

A. Does quantity of nutrients mediates sex specific fitness costs in Callosobruchus maculatus? Małek, D.; **Dańko**, **M. J.;** Czarnołeski, M. In preparation

B. Identifying the Pattern of Human Mortality at Its Front End. Missov, T. I.; Németh, L.; Ribeiro, F; **Dańko, M. J.** Submitted* to Science

C. Density shapes patterns of survival and reproduction in hydromedusa *Eleutheria dichotoma*. Dańko, A.; Schaible, R.; Pijanowska, J.; **Dańko M. J.** Submitted* to Mechanisms of Aging and Development

D. Latitudinal and age-specific patterns of larval mortality in the damselfly Lestes sponsa: Senescence before maturity? **Dańko, M.J.**; Dańko, A.; Golab, M.J.; Stoks R.; Sniegula, S. Submitted to Experimental Gerontology

E. Life history traits are shaped by the interaction of extrinsic mortality and density-dependence. **Dańko, M.J.**; Burger, O.F.; Kozłowski, J. Submitted to PLOS Computational Biology

F. Age-related changes of physiological performance and survivorship of bank voles selected for high aerobic capacity. Rudolf, A.M.; **Dańko, M.J.**; Sadowska, E.T.; Dheyongera, G.; Koteja, P. Submitted to Experimental Gerontology

G. How much can we trust life tables? Sensitivity of mortality measures to right-censoring treatment. Missov, T. I.; Németh, L.; **Dańko, M. J.** Palgrave Communications, 2:15049

And more…..

Selected major research projects of 2016 (chronologically)

Statistical methods used

A. Does quantity of nutrients mediates sex specific fitness costs in Callosobruchus maculatus?
(Frailty parametric proportional hazard models, frailty Cox proportional models, Mortality smoothing, …)

B. Identifying the Pattern of Human Mortality at Its Front End.
(Fitting Gamma-Gompertz-Makeham model via ML, Parametric bootstrap, AIC/Hierarchical LRT models selection,…)

C. Density shapes patterns of survival and reproduction in hydromedusa *Eleutheria dichotoma*.
(Poisson regression, Mortality Smoothing, Fitting Gamma-Gompertz Model via ML,…)

D. Latitudinal and age-specific patterns of larval mortality in the damselfly Lestes sponsa: Senescence before maturity?
(Weighted logrank tests, Fitting family of Gompertz models, AIC/Hierarchical LRT model selection,…)

E. Life history traits are shaped by the interaction of extrinsic mortality and density-dependence.
(Matrix projection models, optimal resource allocations models, Gompertz-Makeham, AFT/PH models,…)

F. Age-related changes of physiological performance and survivorship of bank voles selected for high aerobic capacity. (Weighted logrank tests, Mortality smoothing, Mixture effects models,…)

G. How much can we trust life tables? Sensitivity of mortality measures to right-censoring treatment.
(Fitting Gamma-Gompertz-Makeham model by ML, Parametric bootstrap, Simulation of censoring, LT meaures,…)

Other projects: Lifetables analysis of large databases, PCLM used to approximate LT measures, bootstrap to calculate confidence intervals of LT measurs, weighted linear models, …

Selected major research projects of 2016 (chronologically)

Statistical methods used

A. Does quantity of nutrients mediates sex specific fitness costs in Callosobruchus maculatus?
(Frailty parametric proportional hazard models, frailty Cox proportional models, Mortality smoothing, …)

B. Identifying the Pattern of Human Mortality at Its Front End.
(Fitting Gamma-Gompertz-Makeham model via ML, Parametric bootstrap, AIC/Hierarchical LRT models selection,…)

C. Density shapes patterns of survival and reproduction in hydromedusa *Eleutheria dichotoma*.
(Poisson regression, Mortality Smoothing, Fitting Gamma-Gompertz Model via ML,…)

D. Latitudinal and age-specific patterns of larval mortality in the damselfly Lestes sponsa: Senescence before maturity?
(Weighted logrank tests, Fitting family of Gompertz models, AIC/Hierarchical LRT model selection,…)

E. Life history traits are shaped by the interaction of extrinsic mortality and density-dependence.
(Matrix projection models, optimal resource allocations models, Gompertz-Makeham, AFT/PH models,…)

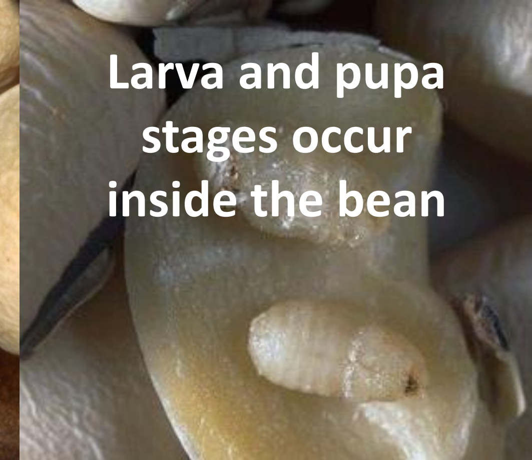
F. Age-related changes of physiological performance and survivorship of bank voles selected for high aerobic capacity. (Weighted logrank tests, Mortality smoothing, Mixture effects models,…)

G. How much can we trust life tables? Sensitivity of mortality measures to right-censoring treatment.
(Fitting Gamma-Gompertz-Makeham model by ML, Parametric bootstrap, Simulation of censoring, LT meaures,…)

Other projects: Lifetables analysis of large databases, PCLM used to approximate LT measures, bootstrap to calculate confidence intervals of LT measurs, weighted linear models, …

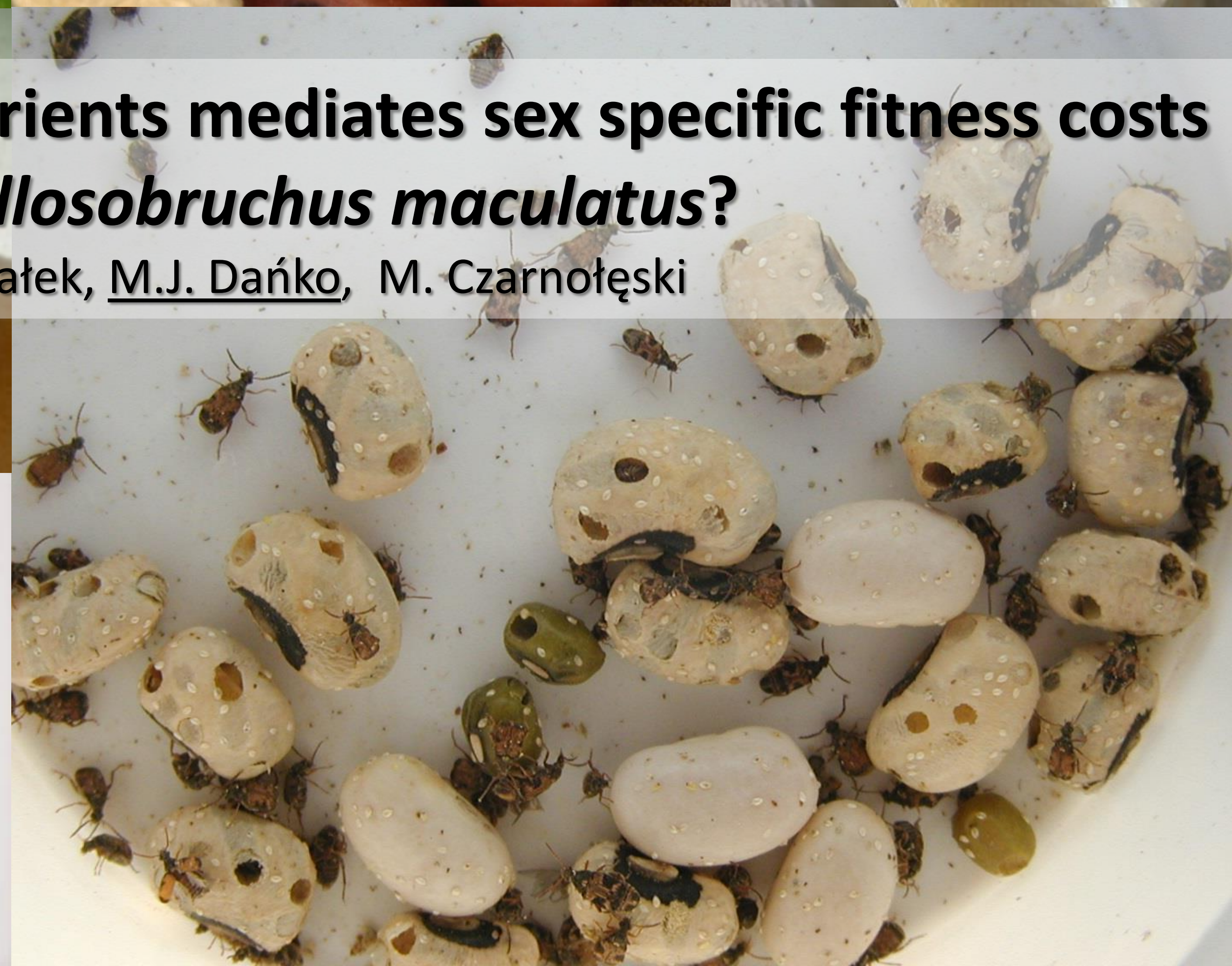
Does quantity of nutrients mediates sex specific fitness costs in *Callosobruchus maculatus*?

D. Małek, M.J. Dańko, M. Czarnołęski



Does quantity of nutrients mediates sex specific fitness costs in *Callosobruchus maculatus*?

D. Małek, M.J. Dańko, M. Czarnołęski



Costs of reproduction

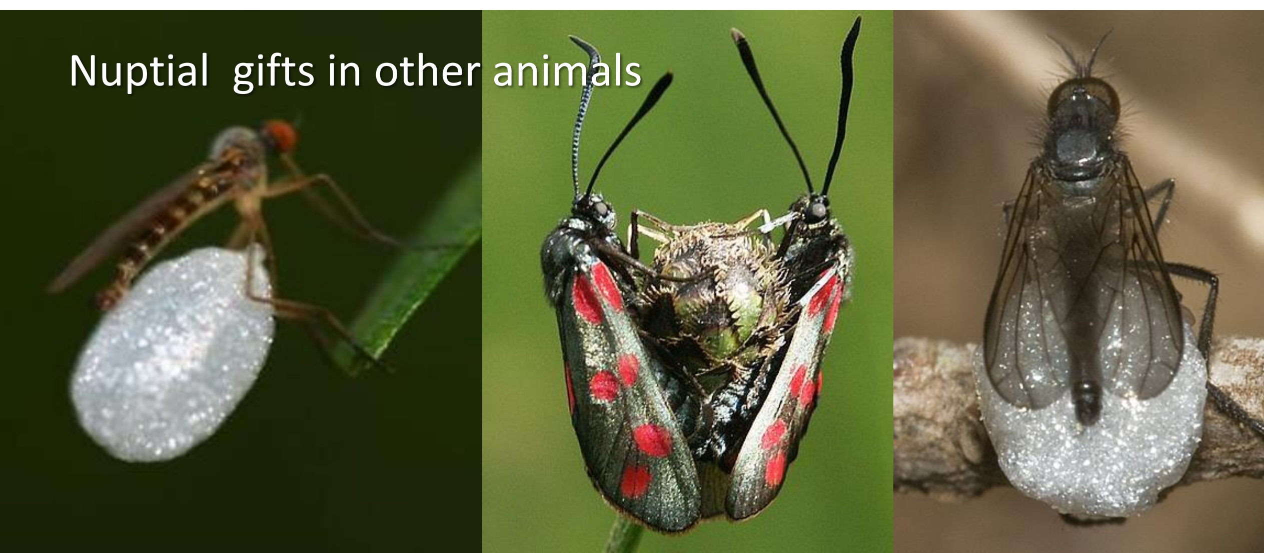
- Two sexes may pay different costs because of different reproductive biology.

Costs of reproduction

- Two sexes may pay different costs because of different reproductive biology.
- Reproductive costs in both sexes can also influence one another, for example by a phenomenon known as **nuptial gifts**.

Costs of reproduction

- Two sexes may pay different costs because of different reproductive biology.
- Reproductive costs in both sexes can also influence one another, for example by a phenomenon known as **nuptial gifts**.



Costs of reproduction

- Two sexes may pay different costs because of different reproductive biology.
- Reproductive costs in both sexes can also influence one another, for example by a phenomenon known as **<u>nuptial gifts</u>**.
- An endogenous gift should be especially costly to a donor because adult individuals of *Callosobruchus maculatus* in laboratory conditions do not ingest food or water, resulting in a very limited energy budget.




Goals of the project (relevant to survival analysis)

Treatment: reproduction allowed, **control:** reproduction not allowed.
Two sexes
Covariates: <u>bean size</u>, <u>adult size</u> and <u>gift size</u> (reproducing animals)
Random effects: mother id

Goals of the project (relevant to survival analysis)

Treatment: reproduction allowed, **control:** reproduction not allowed.
Two sexes
Covariates: <u>bean size</u>, <u>adult size</u> and <u>gift size</u> (reproducing animals)
Random effects: mother id

1) Investigate the effect of sex, bean size, presence of reproduction, and adult size on survival.

2) Investigate the role of nuptial gifts in reproducing males and females for their survival

Frailty models

(Random effects models in survival analysis)

- Frailty term was coined by James Vaupel (Vupel et al. 1979) and describes hidden differences in the survival chances of individuals in a population.

Frailty models
(Random effects models in survival analysis)

- Frailty term was coined by James Vaupel (Vupel et al. 1979) and describes hidden differences in the survival chances of individuals in a population.

- Frailty models:
  - Individual frailty models (each individual has its own frailty)
  - Shared frailty models are random effects model (a group of individuals share the same hidden frailty)

Frailty models

(Random effects models in survival analysis)

- Frailty term was coined by James Vaupel (Vupel et al. 1979) and describes hidden differences in the survival chances of individuals in a population.

- Frailty models:
  - Individual frailty models (each individual has its own frailty)
  - Shared frailty models are <u>random effects model </u>(a group of individuals share the same hidden frailty)

- Shared frailty models:
  - Covered in books, but research is still ongoing
  - Only partially covered in statistical packages

Proportional hazard shared frailty models

Conditional hazard of subject *i* that belongs to the group *g*

$$h_{g,i}(t \mid u_g) = u_g h_0(t) \exp(X_{g,i}^T \beta)$$

Frailty term in group *g*. Realization of a random variable with certain pdf (e.g. gamma distribution)

Baseline hazard (e.g. Weibull)

Vector of regression coefficients applied to a vector of observed covariates for subject *i* in group *g*

Proportional hazard shared frailty models

Conditional hazard of subject *i* that belongs to the group *g*

$$h_{g,i}(t \mid u_g) = u_g h_0(t) \exp(X_{g,i}^T \beta)$$

Frailty term in group *g*. Realization of a random variable with certain pdf (e.g. gamma distribution)

Baseline hazard (e.g. Weibull)

Vector of regression coefficients applied to a vector of observed covariates for subject *i* in group *g*

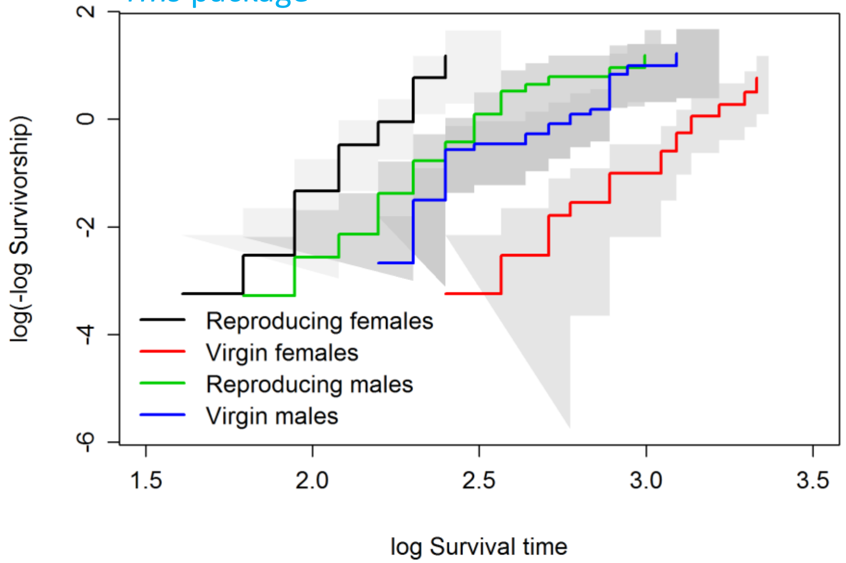
The estimation procedure is based on expectation-maximization (EM) algorithm, which is a sequence of:

A) <u>Expectation</u> ➔ Prediction of $\boldsymbol{u}$ given the estimates of parameters of theoretical frailty distribution, baseline hazard, and regression coefficients.
B) <u>Maximization</u> ➔ Use predicted $\boldsymbol{u}$ to find all parameter estimates by maximization marginal log-likelihood function.

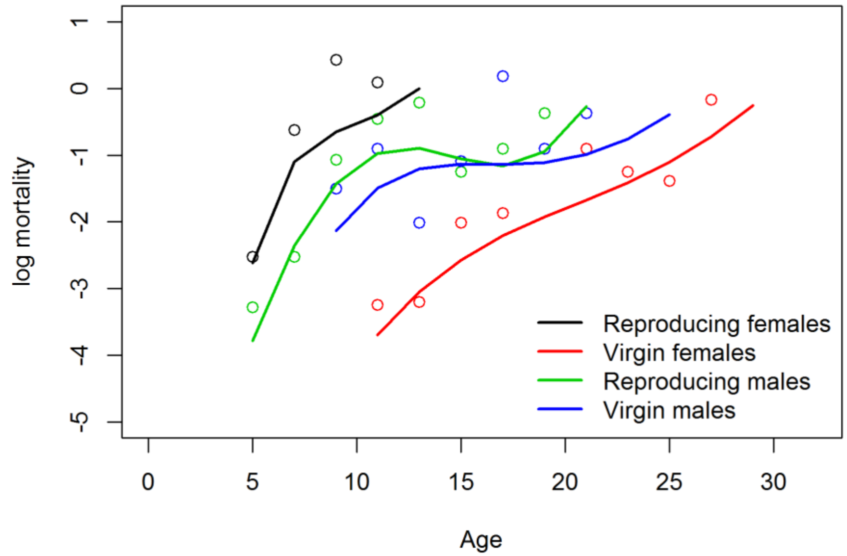
The sequence is repeated until the convergence

Testing proportional hazard assumption
(categorical variables)

rms package

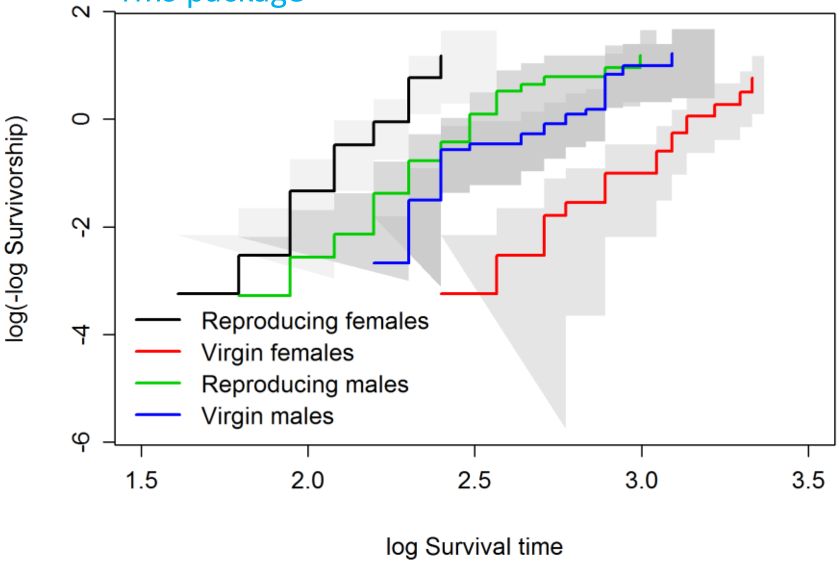


MortalitySmooth package

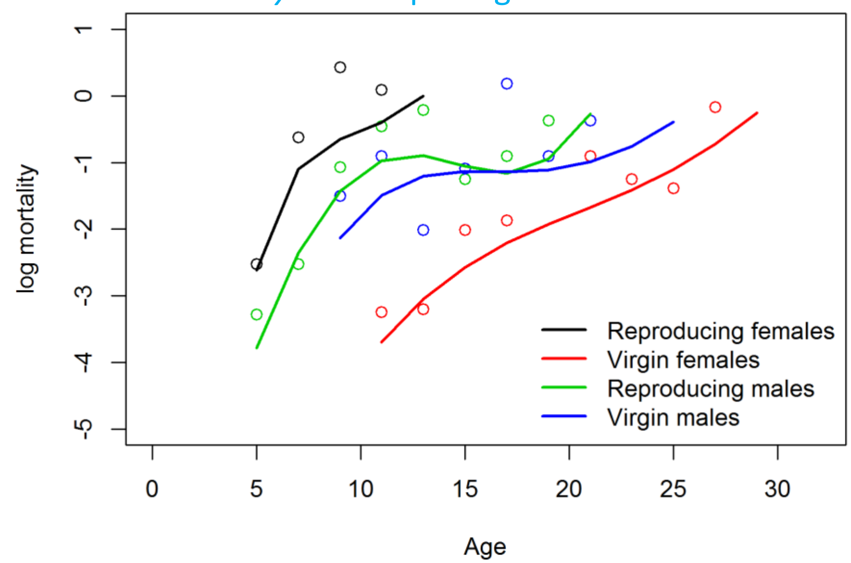


Testing proportional hazard assumption
(categorical variables)

rms package



MortalitySmooth package



Clearly visible interaction between sex and treatment!

Testing proportional hazard assumption
(continuous variables)

- It is hard to perform similar non-parametric analysis on continuous variables
- We will use a semi-parametric approach based on Cox proportional hazard frailty model.

Model: Surv ~ Sex + Treatment + Bean size + Adullt body mass + Sex : Treatment

Testing proportional hazard assumption
(continuous variables)

- It is hard to perform similar non-parametric analysis on continuous variables
- We will use a semi-parametric approach based on Cox proportional hazard frailty model.

Model: Surv ~ Sex + Treatment + Bean size + Adullt body mass + Sex : Treatment

survival package*

	rho	chisq	p
Sex (Males)	-0.0052	0.0035	0.9526
Treatment (Virgin)	0.0677	0.4514	0.5017
Bean size	0.0176	0.0453	0.8315
Adult body mass	0.0139	0.0267	0.8703
Interaction (sex:treatment)	-0.0305	0.1037	0.7474
GLOBAL		0.9675	0.9651

* P. Grambsch and T. Therneau (1994), Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika,* **81**, 515-26.

Testing proportional hazard assumption
(continuous variables)

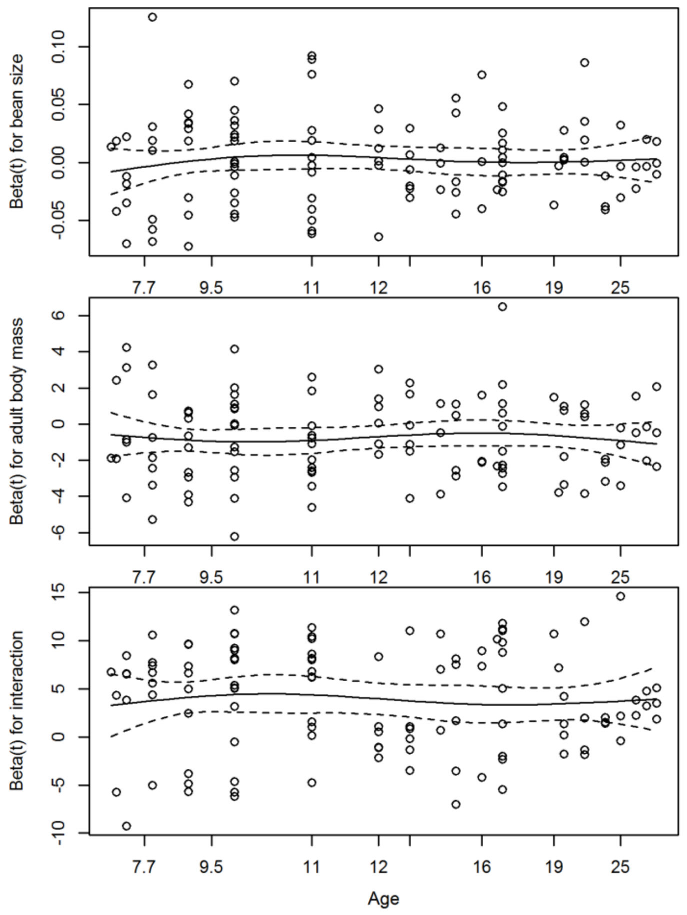
- It is hard to perform similar non-parametric analysis on continuous variables
- We will use a semi-parametric approach based on Cox proportional hazard frailty model.

Model: Surv ~ Sex + Treatment + Bean size + Adullt body mass + Sex : Treatment

survival package*

	rho	chisq	p
Sex (Males)	-0.0052	0.0035	0.9526
Treatment (Virgin)	0.0677	0.4514	0.5017
Bean size	0.0176	0.0453	0.8315
Adult body mass	0.0139	0.0267	0.8703
Interaction (sex:treatment)	-0.0305	0.1037	0.7474
GLOBAL		0.9675	0.9651

Schoenfeld residuals



No clear departures from linearity

* P. Grambsch and T. Therneau (1994), Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika,* **81**, 515-26.

Testing proportional hazard assumption
(continuous variables)

- It is hard to perform similar non-parametric analysis on continuous variables
- We will use a semi-parametric approach based on Cox proportional hazard frailty model.

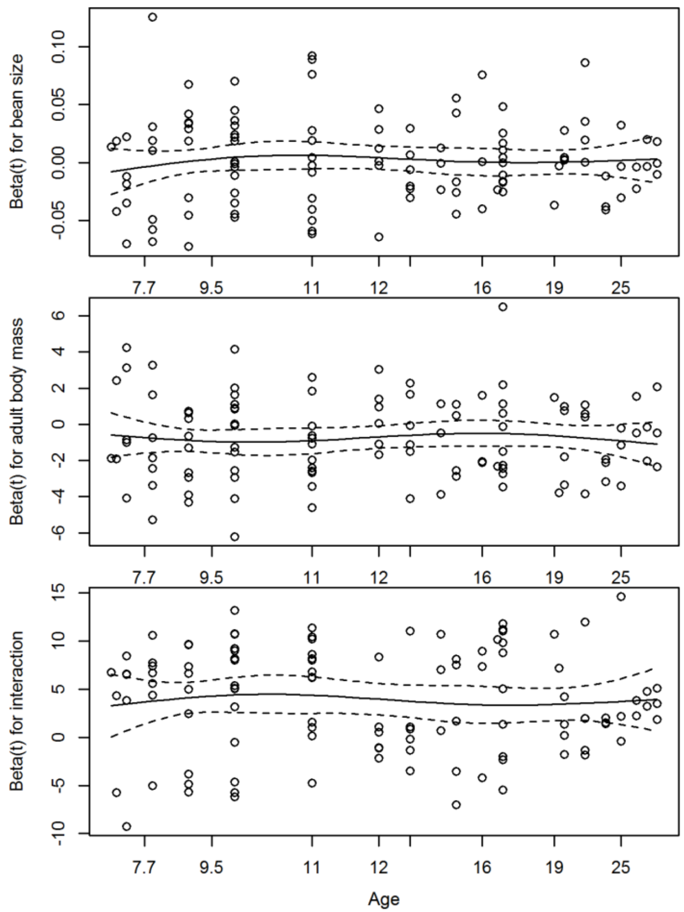
Model: Surv ~ Sex + Treatment + Bean size + Adullt body mass + Sex : Treatment

survival package*

	rho	chisq	p
Sex (Males)	-0.0052	0.0035	0.9526
Treatment (Virgin)	0.0677	0.4514	0.5017
Bean size	0.0176	0.0453	0.8315
Adult body mass	0.0139	0.0267	0.8703
Interaction (sex:treatment)	-0.0305	0.1037	0.7474
GLOBAL		0.9675	0.9651

Conclusion: PH assumptions holds (at least for Cox model).

Schoenfeld residuals



No clear departures from linearity

* P. Grambsch and T. Therneau (1994), Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, **81**, 515-26.

First guess about baseline hazard

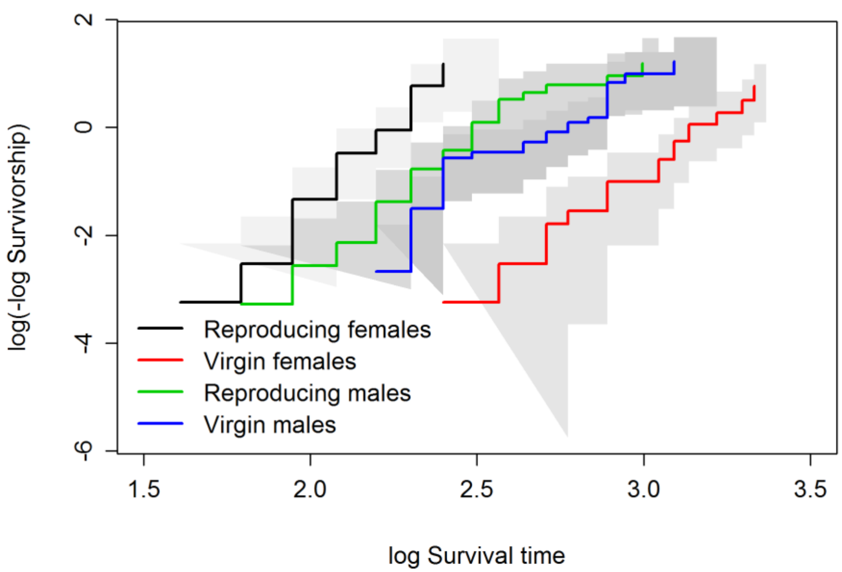
- The plots of log(-log (survivorship)) can help to guess about the baseline
- Log(-log(Survivorship)) = log cumulative hazard

First guess about baseline hazard

- The plots of log(-log (survivorship)) can help to guess about the baseline
- Log(-log(Survivorship)) = log cumulative hazard
- Linear relationship between log cumulative hazard and survival time suggests exponential distribution (constant hazard)
- Linear relationship between log cumulative hazard and log survival time suggests Weibull distribution (Weibull hazard)

First guess about baseline hazard

- The plots of log(-log (survivorship)) can help to guess about the baseline
- Log(-log(Survivorship)) = log cumulative hazard
- Linear relationship between log cumulative hazard and survival time suggests exponential distribution (constant hazard)
- Linear relationship between log cumulative hazard and log survival time suggests Weibull distribution (Weibull hazard)



Curves seems to be linear in early ages

Selection of baseline hazard and frailty distribution

We can use GoF tests, however heterogeneity can affects observed mortality patterns.

A mixture of different Weibull distributions is not necessarily a Weibull distribution

Tests are not decisive, but only informative

Selection of baseline hazard and frailty distribution

We can use GoF tests, however heterogeneity can affects observed mortality patterns.

A mixture of different Weibull distributions is not necessarily a Weibull distribution

Tests are not decisive, but only informative

	p-val	
	KS Bootstrap*	**Generalized Gamma****
Reproducing females	0.6054	0.2486
Virgin females	0.8352	0.5075
Reproducing males	0.2071	0.0219
Virgin males	0.1509	0.1079

- * Parametric bootstrap of null hypothesis of Kolmogorov-Smirnoff statistics

** *EwGOF* package

* Jogesh Babu, G., and C. R. Rao. Goodness-of-fit tests when parameters are estimated. Sankhya: The Indian Journal of Statistics 66 (2004): 63-74.
** Krit M., Gaudoin O., Xie M. and Remy E., Simplified likelihood goodness-of-fit tests for the Weibull distribution, *Communications in Statistics - Simulation and Computation*.

Selection of baseline hazard and frailty distribution

We can use GoF tests, however heterogeneity can affects observed mortality patterns.

A mixture of different Weibull distributions is not necessarily a Weibull distribution

Tests are not decisive, but only informative

	p-val	
	KS Bootstrap*	**Generalized Gamma****
Reproducing females	0.6054	0.2486
Virgin females	0.8352	0.5075
Reproducing males	0.2071	0.0219
Virgin males	0.1509	0.1079

- Parametric bootstrap of null hypothesis of Kolmogorov-Smirnoff statistics

** *EwGOF* package

* Jogesh Babu, G., and C. R. Rao. Goodness-of-fit tests when parameters are estimated. Sankhya: The Indian Journal of Statistics 66 (2004): 63-74.
** Krit M., Gaudoin O., Xie M. and Remy E., Simplified likelihood goodness-of-fit tests for the Weibull distribution, *Communications in Statistics - Simulation and Computation*.

Selection of baseline hazard and frailty distribution

We can use GoF tests, however heterogeneity can affects observed mortality patterns.

A mixture of different Weibull distributions is not necessarily a Weibull distribution

Tests are not decisive, but only informative

fitdistrplus package

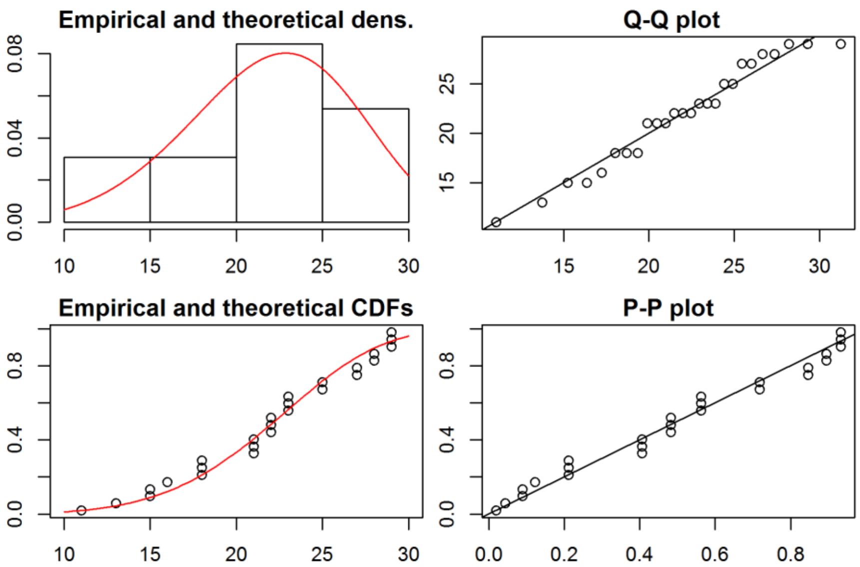


FIG. Empirical vs. Theoretical Weibull distribution for **virgin females**.

	p-val	
	KS Bootstrap*	Generalized Gamma**
Reproducing females	0.6054	0.2486
Virgin females	0.8352	0.5075
Reproducing males	0.2071	0.0219
Virgin males	0.1509	0.1079

- Parametric bootstrap of null hypothesis of Kolmogorov-Smirnoff statistics

** *EwGOF* package

* Jogesh Babu, G., and C. R. Rao. Goodness-of-fit tests when parameters are estimated. Sankhya: The Indian Journal of Statistics 66 (2004): 63-74.
** Krit M., Gaudoin O., Xie M. and Remy E., Simplified likelihood goodness-of-fit tests for the Weibull distribution, *Communications in Statistics - Simulation and Computation*.

Selection of baseline hazard and frailty distribution

We can use GoF tests, however heterogeneity can affects observed mortality patterns.

A mixture of different Weibull distributions is not necessarily a Weibull distribution

Tests are not decisive, but only informative

fitdistrplus package

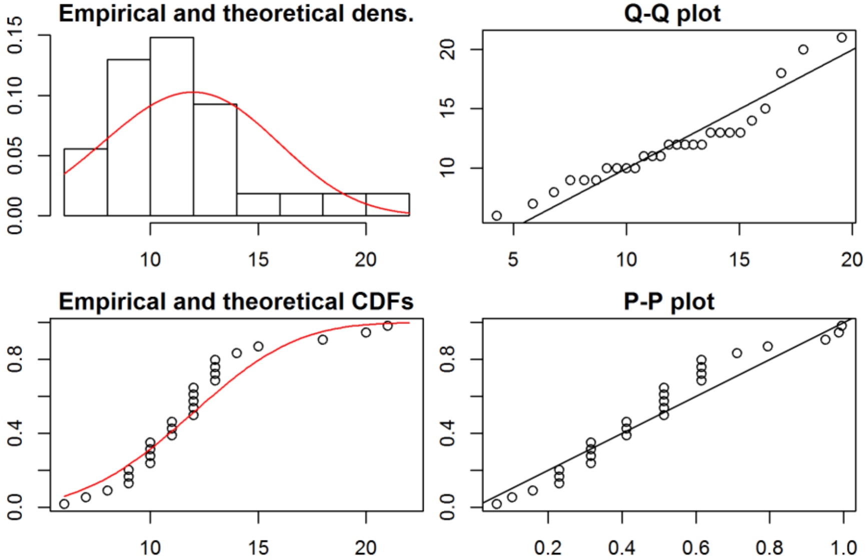


FIG. Empirical vs. Theoretical Weibull distribution for **reproducing males**.

	p-val	
	KS Bootstrap*	Generalized Gamma**
Reproducing females	0.6054	0.2486
Virgin females	0.8352	0.5075
Reproducing males	0.2071	0.0219
Virgin males	0.1509	0.1079

- Parametric bootstrap of null hypothesis of Kolmogorov-Smirnoff statistics

** *EwGOF* package

* Jogesh Babu, G., and C. R. Rao. Goodness-of-fit tests when parameters are estimated. Sankhya: The Indian Journal of Statistics 66 (2004): 63-74.
** Krit M., Gaudoin O., Xie M. and Remy E., Simplified likelihood goodness-of-fit tests for the Weibull distribution, *Communications in Statistics - Simulation and Computation*.

Selection of baseline hazard and frailty distribution

We can use GoF tests, however heterogeneity can affects observed mortality patterns.

A mixture of different Weibull distributions is not necessarily a Weibull distribution

Tests are not decisive, but only informative

fitdistrplus package

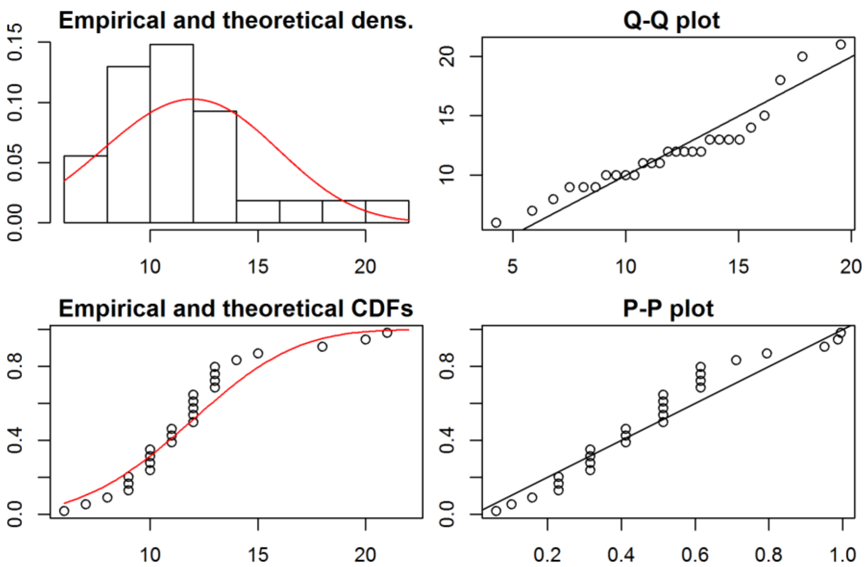


FIG. Empirical vs. Theoretical Weibull distribution for **reproducing males**.

	p-val	
	KS Bootstrap*	Generalized Gamma**
Reproducing females	0.6054	0.2486
Virgin females	0.8352	0.5075
Reproducing males	0.2071	0.0219
Virgin males	0.1509	0.1079

- Parametric bootstrap of null hypothesis of Kolmogorov-Smirnoff statistics

** *EwGOF* package

* Jogesh Babu, G., and C. R. Rao. Goodness-of-fit tests when parameters are estimated. Sankhya: The Indian Journal of Statistics 66 (2004): 63-74.
** Krit M., Gaudoin O., Xie M. and Remy E., Simplified likelihood goodness-of-fit tests for the Weibull distribution, *Communications in Statistics - Simulation and Computation*.

Model selection based on AIC

Model: Surv ~ Sex + Treatment + Bean size + Adult body mass + Sex : Treatment

	Frailty distribution	
Baseline hazard	none	gamma
Gompertz	616.19	613.87
Weibull	584.98	583.68
Exponential	795.7	797.77

Selection of baseline hazard and frailty distribution

We can use GoF tests, however heterogeneity can affects observed mortality patterns.

A mixture of different Weibull distributions is not necessarily a Weibull distribution

Tests are not decisive, but only informative

fitdistrplus package

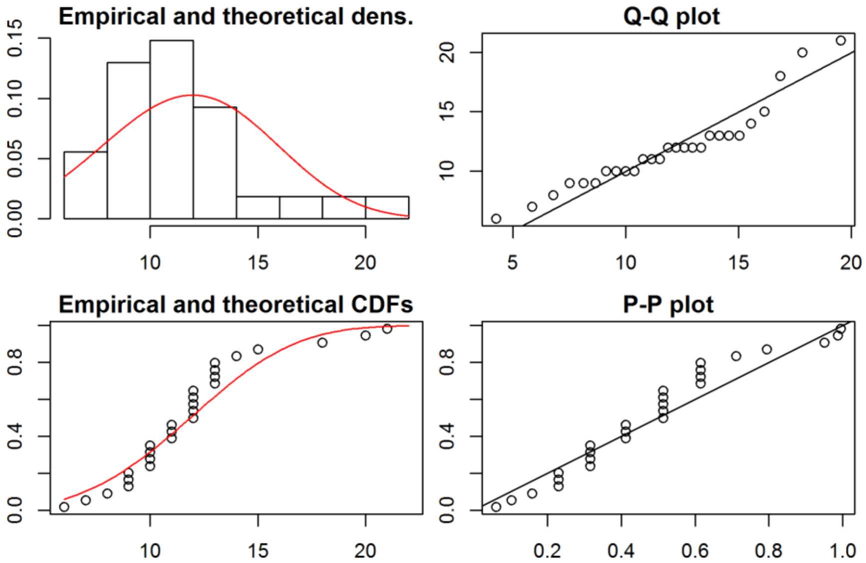


FIG. Empirical vs. Theoretical Weibull distribution for **reproducing males**.

	p-val	
	KS Bootstrap*	Generalized Gamma**
Reproducing females	0.6054	0.2486
Virgin females	0.8352	0.5075
Reproducing males	0.2071	0.0219
Virgin males	0.1509	0.1079

- Parametric bootstrap of null hypothesis of Kolmogorov-Smirnoff statistics

** *EwGOF* package

* Jogesh Babu, G., and C. R. Rao. Goodness-of-fit tests when parameters are estimated. Sankhya: The Indian Journal of Statistics 66 (2004): 63-74.
** Krit M., Gaudoin O., Xie M. and Remy E., Simplified likelihood goodness-of-fit tests for the Weibull distribution, *Communications in Statistics - Simulation and Computation*.

Model selection based on AIC

Model: Surv ~ Sex + Treatment + Bean size + Adult body mass + Sex : Treatment

	Frailty distribution	
Baseline hazard	none	gamma
Gompertz	616.19	613.87
Weibull	584.98	583.68
Exponential	795.7	797.77

Basic model and model selection

Improved maximization part of EM algorithm of *parfm* package

	ESTIMATE	SE	p-val
Theta (frailty par.)	0.1572	0.1290	
Rho (Weibull shape par.)	4.7296	0.3877	
Lambda (Weibull scale par.)	0.0026	0.0036	
Sex (Males)	-2.8240	0.4835	0.0000
Treatment (Virgin)	-4.5469	0.4850	0.0000
Bean size	0.0011	0.0031	0.7263
Addult body mass	-0.6894	0.1998	0.0006
Interaction (sex:treatment)	3.8895	0.5384	0.0000

Basic model with included Sex:Treatment interaction.
The p-values are calculated from Wald test.

Basic model and model selection

Improved maximization part of EM algorithm
of *parfm* package

	ESTIMATE	SE	p-val
Theta (frailty par.)	0.1572	0.1290	
Rho (Weibull shape par.)	4.7296	0.3877	
Lambda (Weibull scale par.)	0.0026	0.0036	
Sex (Males)	-2.8240	0.4835	0.0000
Treatment (Virgin)	-4.5469	0.4850	0.0000
Bean size	0.0011	0.0031	0.7263
Addult body mass	-0.6894	0.1998	0.0006
Interaction (sex:treatment)	3.8895	0.5384	0.0000

Basic model with included Sex:Treatment interaction.
The p-values are calculated from Wald test.

We will perform model selection by sequentially adding single 2-way or 3-way interaction and performing Likelihood ratio test.

The estimates of the basic model will be used as starting values for optimization method of more complicated models.

Basic model and model selection

Improved maximization part of EM algorithm of *parfm* package

	ESTIMATE	SE	p-val
Theta (frailty par.)	0.1572	0.1290	
Rho (Weibull shape par.)	4.7296	0.3877	
Lambda (Weibull scale par.)	0.0026	0.0036	
Sex (Males)	-2.8240	0.4835	0.0000
Treatment (Virgin)	-4.5469	0.4850	0.0000
Bean size	0.0011	0.0031	0.7263
Addult body mass	-0.6894	0.1998	0.0006
Interaction (sex:treatment)	3.8895	0.5384	0.0000

Basic model with included Sex:Treatment interaction.
The p-values are calculated from Wald test.

We will perform model selection by sequentially adding single 2-way or 3-way interaction and performing Likelihood ratio test.

The estimates of the basic model will be used as starting values for optimization method of more complicated models.

LRT added to *parfm* package

Added interaction	chisq	df	pval
Bean size : adult body mass	2.22	1	0.1362
Sex : adult body mass	0.27	1	0.6003
Sex : bean size	1.51	1	0.2195
Treatment : adult body mass	1.91	1	0.1673
Treatment : bean size	1.75	1	0.1858
Sex : treatment : bean size *	2.36	4	0.6697
Sex : treatment : adult body mass *	2.46	4	0.6519
Sex : bean size : adult body mass	2.25	2	0.3243
Treatment : bean size : adult body mass	1.76	2	0.4156

* ”Negative variance” problem

The model cannot be further improved by ”sequential” LRT

Basic model and model selection

Improved maximization part of EM algorithm of *parfm* package

	ESTIMATE	SE	p-val
Theta (frailty par.)	0.1572	0.1290	
Rho (Weibull shape par.)	4.7296	0.3877	
Lambda (Weibull scale par.)	0.0026	0.0036	
Sex (Males)	-2.8240	0.4835	0.0000
Treatment (Virgin)	-4.5469	0.4850	0.0000
Bean size	0.0011	0.0031	0.7263
Addult body mass	-0.6894	0.1998	0.0006
Interaction (sex:treatment)	3.8895	0.5384	0.0000

Basic model with included Sex:Treatment interaction. The p-values are calculated from Wald test.

We will perform model selection by sequentially adding single 2-way or 3-way interaction and performing Likelihood ratio test.

The estimates of the basic model will be used as starting values for optimization method of more complicated models.

Variance inflation factors (VIF) measure how much the variance of the estimated regression coefficients are inflated as compared to when the predictor variables are not linearly related.

LRT added to *parfm* package

Added interaction	chisq	df	pval
Bean size : adult body mass	2.22	1	0.1362
Sex : adult body mass	0.27	1	0.6003
Sex : bean size	1.51	1	0.2195
Treatment : adult body mass	1.91	1	0.1673
Treatment : bean size	1.75	1	0.1858
Sex : treatment : bean size	2.36	4	0.6697
Sex : treatment : adult body mass *	2.46	4	0.6519
Sex : bean size : adult body mass	2.25	2	0.3243
Treatment : bean size : adult body mass	1.76	2	0.4156

* ”Negative variance” problem

The model cannot be further improved by ”sequential” LRT

vif() of *rms* package adopted to work with *parfm* package

Testing collinearity by Variance Inflation Factor (VIF)

Rule of thumb: VIF<10

	VIF
Theta (frailty par.)	1.30
Rho (Weibull shape par.)	3.14
Sex (Males)	5.40
Treatment (Virgin)	5.62
Bean size	1.12
Addult body mass	2.69
Interaction (sex:treatment)	4.75

Predictions of the model – marginal hazard

My new extension to *parfm* package

Estimated hazard of subject i

Estimated survivorship of subject i

$$\bar{\mu}(x) = \frac{\sum_i h_i\left(x \mid u_i, X_i, \hat{\beta}, \hat{\rho}, \hat{\lambda}\right)\exp\left[-H_i\left(x \mid u_i, X_i, \hat{\beta}, \hat{\rho}, \hat{\lambda}\right)\right]}{\sum_i \exp\left[-H_i\left(x \mid u_i, X_i, \hat{\beta}, \hat{\rho}, \hat{\lambda}\right)\right]}$$

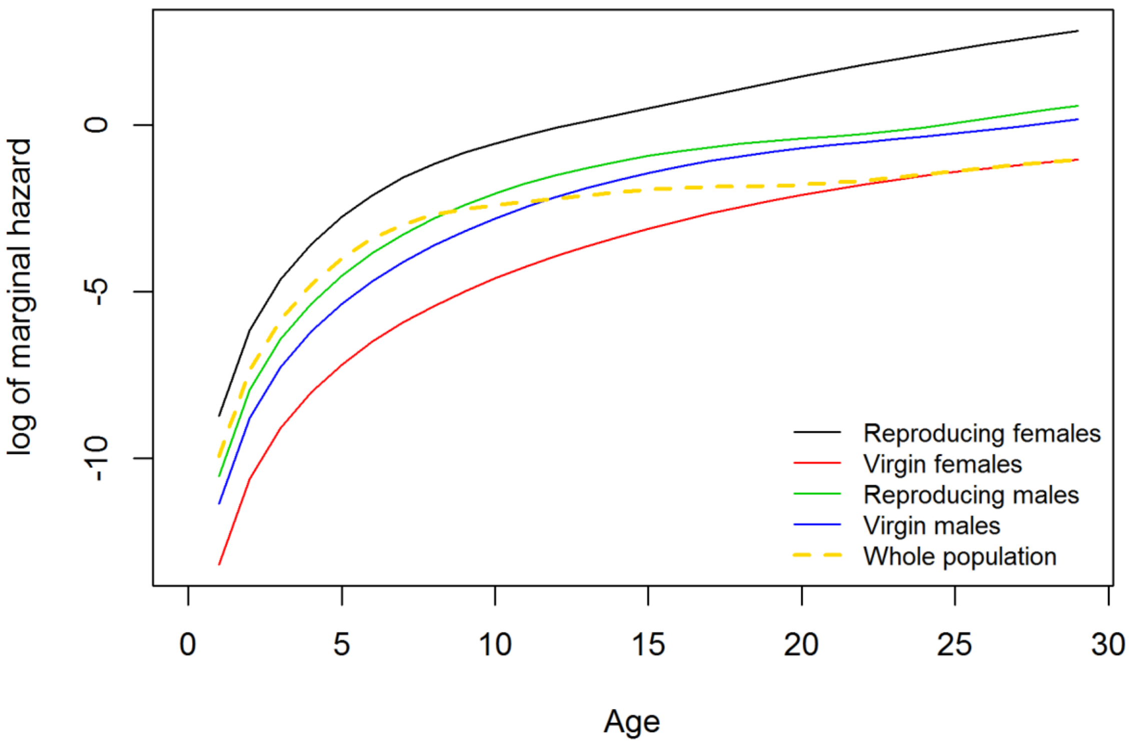
Predictions of the model – marginal hazard

My new extension to *parfm* package

Estimated hazard of subject i

Estimated survivorship of subject i

$$\bar{\mu}(x) = \frac{\sum_i h_i\left(x \mid u_i, X_i, \hat{\beta}, \hat{\rho}, \hat{\lambda}\right)\exp\left[-H_i\left(x \mid u_i, X_i, \hat{\beta}, \hat{\rho}, \hat{\lambda}\right)\right]}{\sum_i \exp\left[-H_i\left(x \mid u_i, X_i, \hat{\beta}, \hat{\rho}, \hat{\lambda}\right)\right]}$$



Marginal hazard for each sex and treatment calculated with respect to individual differences in frailty, adult body mass, and bean size

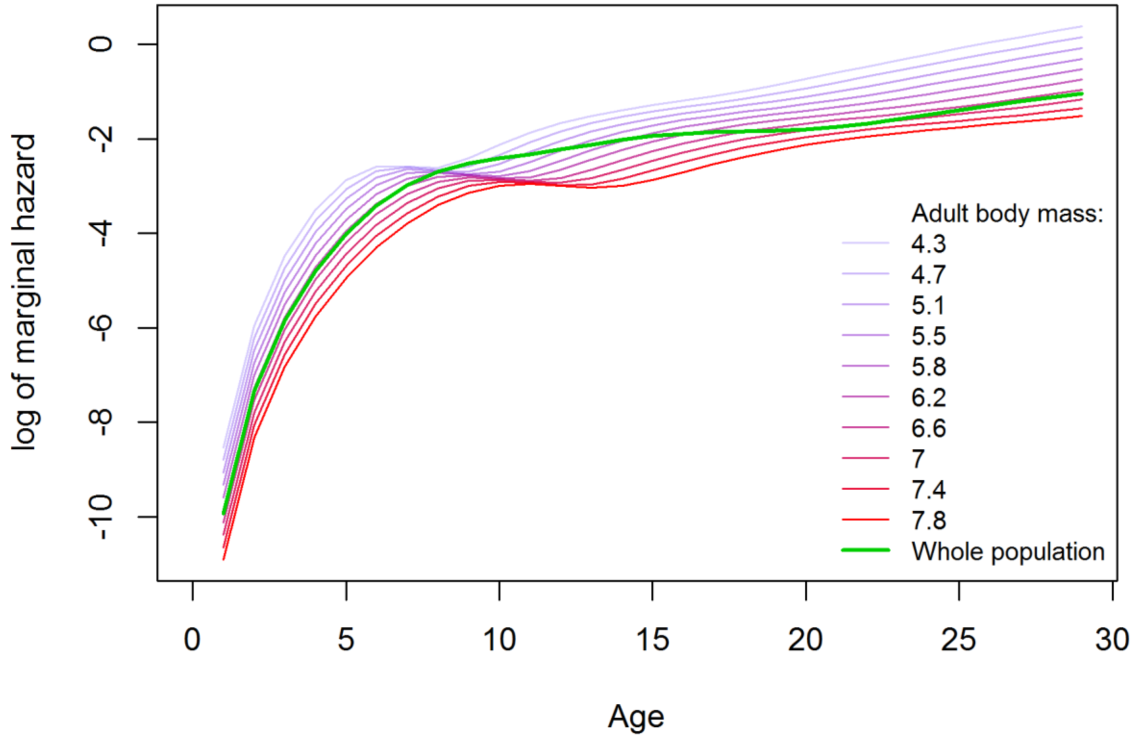
Predictions of the model – marginal hazard

My new extension to *parfm* package

Estimated hazard of subject i

Estimated survivorship of subject i

$$\bar{\mu}(x) = \frac{\sum_i h_i\left(x \mid u_i, X_i, \hat{\beta}, \hat{\rho}, \hat{\lambda}\right)\exp\left[-H_i\left(x \mid u_i, X_i, \hat{\beta}, \hat{\rho}, \hat{\lambda}\right)\right]}{\sum_i \exp\left[-H_i\left(x \mid u_i, X_i, \hat{\beta}, \hat{\rho}, \hat{\lambda}\right)\right]}$$



Marginal hazard as a function of adult body mass calculated with respect to individual differences in frailty, sex, treatment, and bean size

Analyzing model fit by Martingale residuals

My new extension to *parfm* package

Frailty of subject *i*

Follow-up (event) time of subject *i*

$$\hat{M}_i = \delta_i - u_i \hat{H}_0(t_i) \exp\left(\hat{\beta}_1 X_{1,i} + \ldots + \hat{\beta}_n X_{n,i}\right)$$

Event indicator
(0=censoring, 1=death)

Estimated cumulative
baseline hazard

Estimated coefficients applied to observed
covariate for subject *i*

Analyzing model fit by Martingale residuals

My new extension to *parfm* package

Frailty of subject i

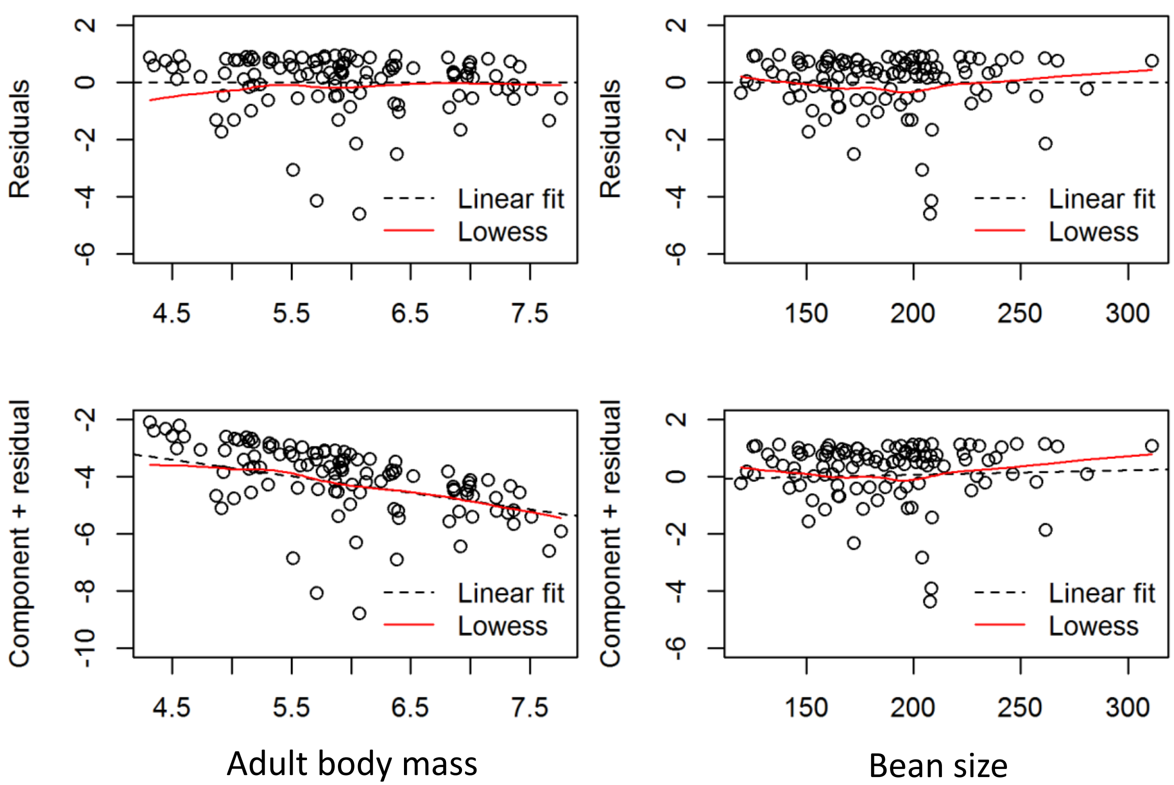
Follow-up (event) time of subject i

$$\hat{M}_i = \delta_i - u_i \hat{H}_0(t_i) \exp\left(\hat{\beta}_1 X_{1,i} + \ldots + \hat{\beta}_n X_{n,i}\right)$$

Event indicator
(0=censoring, 1=death)

Estimated cumulative
baseline hazard

Estimated coefficients applied to observed
covariate for subject i



Modeling the effect of nuptial gift size

Only reproducing animals are relevant

Model will include new set of independent variables:

- Sex
- Gift size
- Bean size
- Adult body mass

Basic model: Surv ~ Sex + Gift size + Bean size + Adult body mass

Modeling the effect of nuptial gift size

Only reproducing animals are relevant

Model will include new set of independent variables:

- Sex
- Gift size
- Bean size
- Adult body mass

Basic model: Surv ~ Sex + Gift size + Bean size + Adult body mass

It is hard to guess about interactions as we cannot easily plot them.

We cannot include all → Building too complicated models can:

- Greatly increase VIF
- Can lead to lack of convergence in EM algorithm
- Make computations extremely long

From the same reasons: exhaustive model selection (e.g. AIC) would be ineffective.

Gift size

Model selection via Hierarchical Likelihood Ratio Test (hLRT)

My new extension to *parfm* package

Model without interaction, but all main effects are present

Basic model: Surv ~ Sex + Gift size + Bean size + Adullt body mass

#1 Sequential LRT test of each of single interaction terms

	logLikelihood	chisq	df	pval
Sex : gift size	-119.77	7.9216	1	0.0049
Sex : bean size	-123.69	0.0886	1	0.7660
Sex : Adult body mass	-123.55	0.3616	1	0.5476
Gift size : bean size	-123.68	0.0932	1	0.7602
Gift size : Adult body mass	-120.56	6.3355	1	0.0118
Bean size : Adult body mass	-123.43	0.6088	1	0.4353

Gift size

Model selection via Hierarchical Likelihood Ratio Test (hLRT)

My new extension to *parfm* package

Model without interaction, but all main effects are present

Basic model: Surv ~ Sex + Gift size + Bean size + Adullt body mass

#1 Sequential LRT test of each of single interaction terms

	logLikelihood	chisq	df	pval
Sex : gift size	-119.77	7.9216	1	0.0049
Sex : bean size	-123.69	0.0886	1	0.7660
Sex : Adult body mass	-123.55	0.3616	1	0.5476
Gift size : bean size	-123.68	0.0932	1	0.7602
Gift size : Adult body mass	-120.56	6.3355	1	0.0118
Bean size : Adult body mass	-123.43	0.6088	1	0.4353

New model = Basic model + Sex : gift size

#2 Selection of significant interactionthat that has the highest log likelihood, then proceed as in #1

Gift size

Model selection via Hierarchical Likelihood Ratio Test (hLRT)

My new extension to *parfm* package

Model without interaction, but all main effects are present

Basic model: Surv ~ Sex + Gift size + Bean size + Adullt body mass

#1 Sequential LRT test of each of single interaction terms

	logLikelihood	chisq	df	pval
Sex : gift size	-119.77	7.9216	1	0.0049
Sex : bean size	-123.69	0.0886	1	0.7660
Sex : Adult body mass	-123.55	0.3616	1	0.5476
Gift size : bean size	-123.68	0.0932	1	0.7602
Gift size : Adult body mass	-120.56	6.3355	1	0.0118
Bean size : Adult body mass	-123.43	0.6088	1	0.4353

New model = Basic model + Sex : gift size

#2 Selection of significant interactionthat that has the highest log likelihood, then proceed as in #1

	logLikelihood	chisq	df	pval
Sex : bean size	-119.43	0.6826	1	0.4087
Sex : Adult body mass	-119.56	0.4282	1	0.5129
Gift size : bean size	-119.66	0.2215	1	0.6379
Gift size : Adult body mass	-119.46	0.6210	1	0.4307
Bean size : Adult body mass	-118.30	2.9423	1	0.0863

Gift size

Model selection via Hierarchical Likelihood Ratio Test (hLRT)

My new extension to *parfm* package

Model without interaction, but all main effects are present

Basic model: Surv ~ Sex + Gift size + Bean size + Adullt body mass

#1 Sequential LRT test of each of single interaction terms

	logLikelihood	chisq	df	pval
Sex : gift size	-119.77	7.9216	1	0.0049
Sex : bean size	-123.69	0.0886	1	0.7660
Sex : Adult body mass	-123.55	0.3616	1	0.5476
Gift size : bean size	-123.68	0.0932	1	0.7602
Gift size : Adult body mass	-120.56	6.3355	1	0.0118
Bean size : Adult body mass	-123.43	0.6088	1	0.4353

New model = Basic model + Sex : gift size

#2 Selection of significant interactionthat that has the highest log likelihood, then proceed as in #1

	logLikelihood	chisq	df	pval
Sex : bean size	-119.43	0.6826	1	0.4087
Sex : Adult body mass	-119.56	0.4282	1	0.5129
Gift size : bean size	-119.66	0.2215	1	0.6379
Gift size : Adult body mass	-119.46	0.6210	1	0.4307
Bean size : Adult body mass	-118.30	2.9423	1	0.0863

No further improvement is possible.

Gift size

Model selection via Hierarchical Likelihood Ratio Test (hLRT)

My new extension to *parfm* package

Model without interaction, but all main effects are present

Basic model: Surv ~ Sex + Gift size + Bean size + Adullt body mass

#1 Sequential LRT test of each of single interaction terms

	logLikelihood	chisq	df	pval
Sex : gift size	-119.77	7.9216	1	0.0049
Sex : bean size	-123.69	0.0886	1	0.7660
Sex : Adult body mass	-123.55	0.3616	1	0.5476
Gift size : bean size	-123.68	0.0932	1	0.7602
Gift size : Adult body mass	-120.56	6.3355	1	0.0118
Bean size : Adult body mass	-123.43	0.6088	1	0.4353

New model = Basic model + Sex : gift size

#2 Selection of significant interactionthat that has the highest log likelihood, then proceed as in #1

	logLikelihood	chisq	df	pval
Sex : bean size	-119.43	0.6826	1	0.4087
Sex : Adult body mass	-119.56	0.4282	1	0.5129
Gift size : bean size	-119.66	0.2215	1	0.6379
Gift size : Adult body mass	-119.46	0.6210	1	0.4307
Bean size : Adult body mass	-118.30	2.9423	1	0.0863

No further improvement is possible.

	ESTIMATE	SE	p-val
Theta (frailty par.)	0.2558	0.2314	
Rho (Weibull shape par.)	5.4213	0.7059	
Lambda (Weibull scale par.)	0.0009	0.0021	
Sex (Males)	-4.8499	1.0594	0.0000
Bean size	-4.2520	2.4488	0.0825
Addult body mass	-0.0045	0.0051	0.3764
Gift size	-0.4326	0.2837	0.1272
Interaction (sex : gift size)	8.5111	3.1087	0.0062

Gift size

Model selection via Hierarchical Likelihood Ratio Test (hLRT)

My new extension to *parfm* package

Model without interaction, but all main effects are present

Basic model: Surv ~ Sex + Gift size + Bean size + Adullt body mass

#1 Sequential LRT test of each of single interaction terms

	logLikelihood	chisq	df	pval
Sex : gift size	-119.77	7.9216	1	0.0049
Sex : bean size	-123.69	0.0886	1	0.7660
Sex : Adult body mass	-123.55	0.3616	1	0.5476
Gift size : bean size	-123.68	0.0932	1	0.7602
Gift size : Adult body mass	-120.56	6.3355	1	0.0118
Bean size : Adult body mass	-123.43	0.6088	1	0.4353

New model = Basic model + Sex : gift size

#2 Selection of significant interactionthat that has the highest log likelihood, then proceed as in #1

	logLikelihood	chisq	df	pval
Sex : bean size	-119.43	0.6826	1	0.4087
Sex : Adult body mass	-119.56	0.4282	1	0.5129
Gift size : bean size	-119.66	0.2215	1	0.6379
Gift size : Adult body mass	-119.46	0.6210	1	0.4307
Bean size : Adult body mass	-118.30	2.9423	1	0.0863

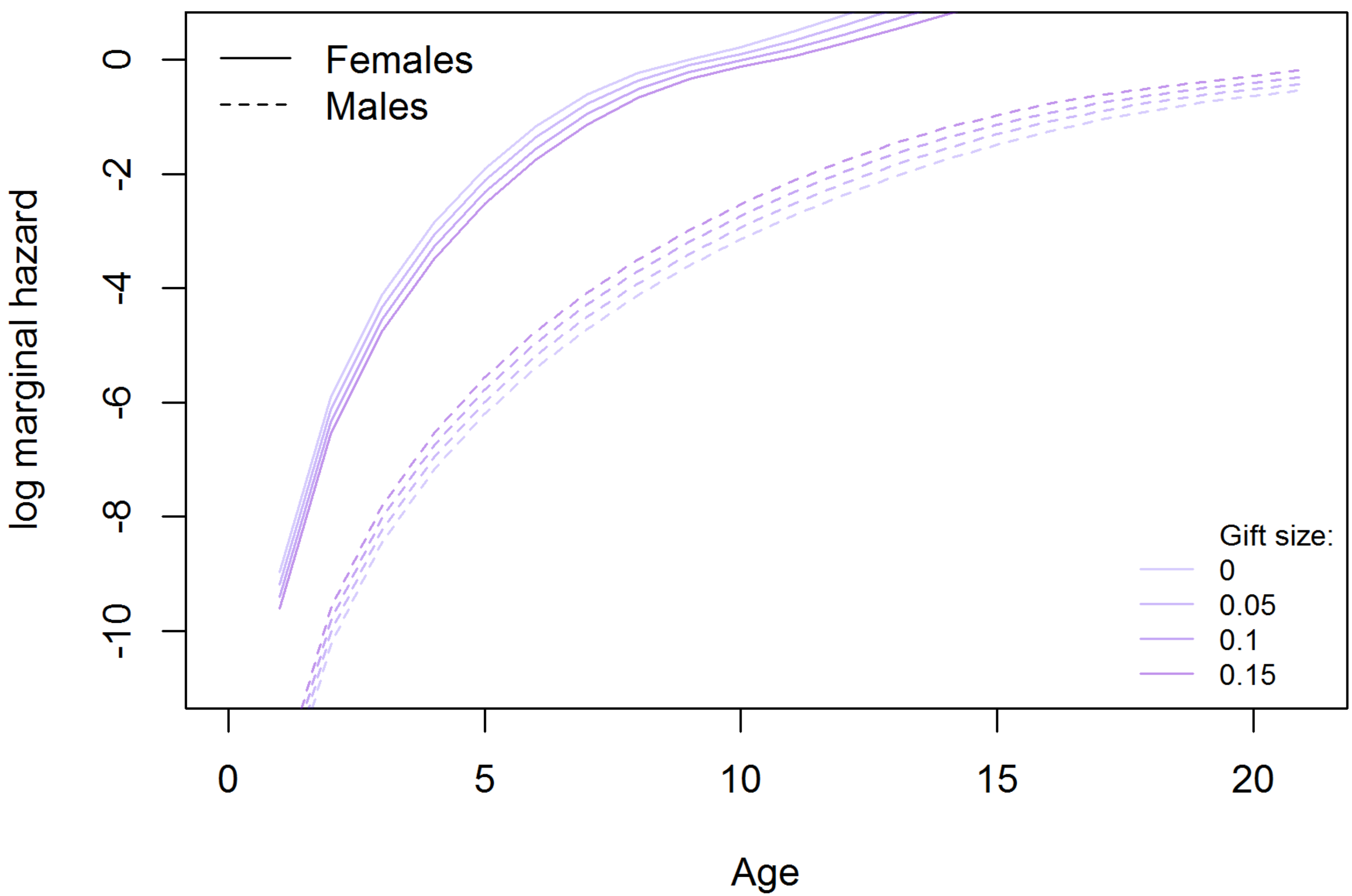
No further improvement is possible.

	ESTIMATE	SE	p-val
Theta (frailty par.)	0.2558	0.2314	
Rho (Weibull shape par.)	5.4213	0.7059	
Lambda (Weibull scale par.)	0.0009	0.0021	
Sex (Males)	-4.8499	1.0594	0.0000
Bean size	-4.2520	2.4488	0.0825
Addult body mass	-0.0045	0.0051	0.3764
Gift size	-0.4326	0.2837	0.1272
Interaction (sex : gift size)	8.5111	3.1087	0.0062

Non-significant terms (like bean size or adult mass) could be dropped from the model (LRT) to decrease VIF (some VIFs slightly exceeded 10, not shown).

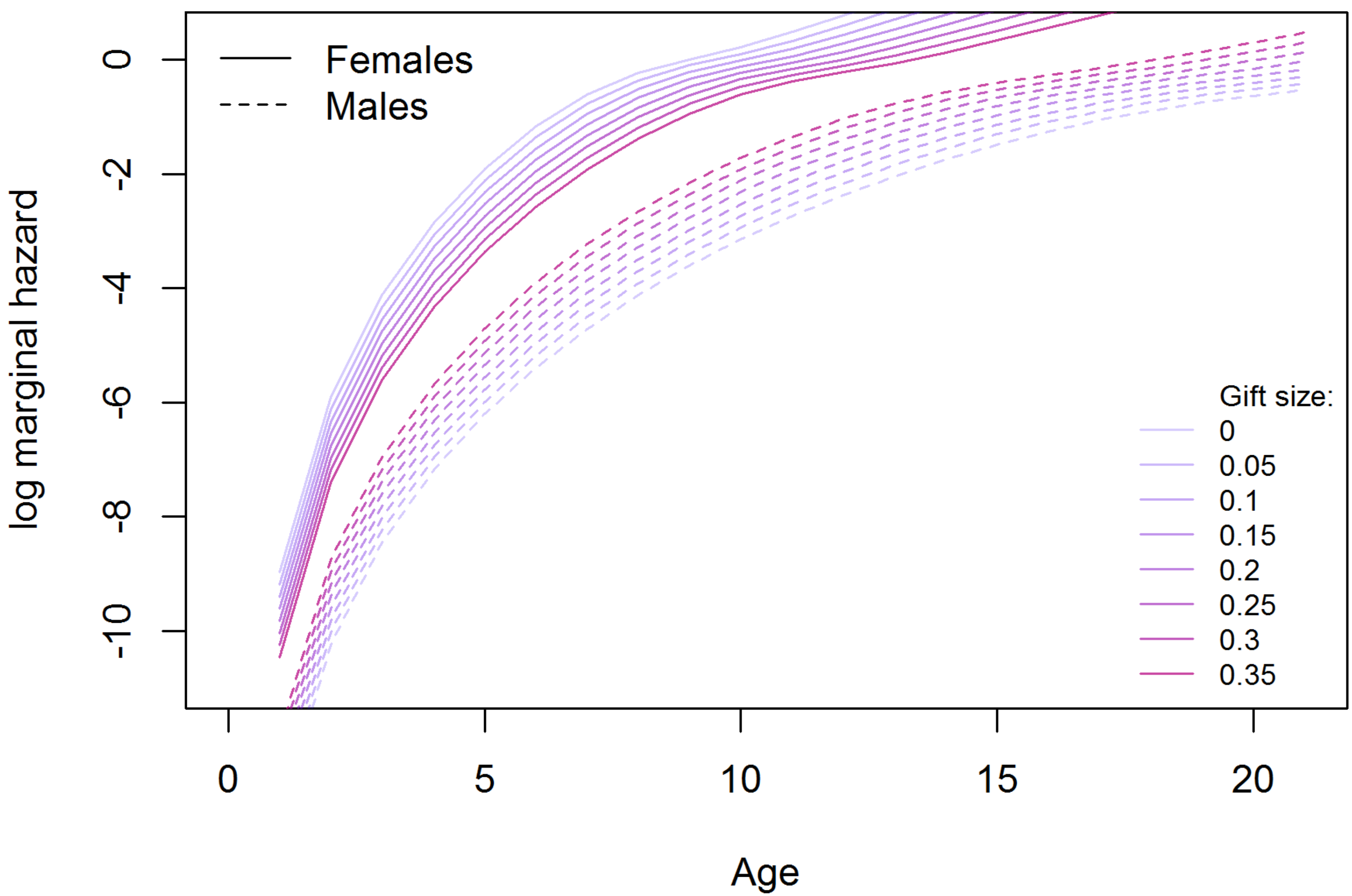
Effect of gift size on marginal hazard rate in males and females

My new extension to *parfm* package



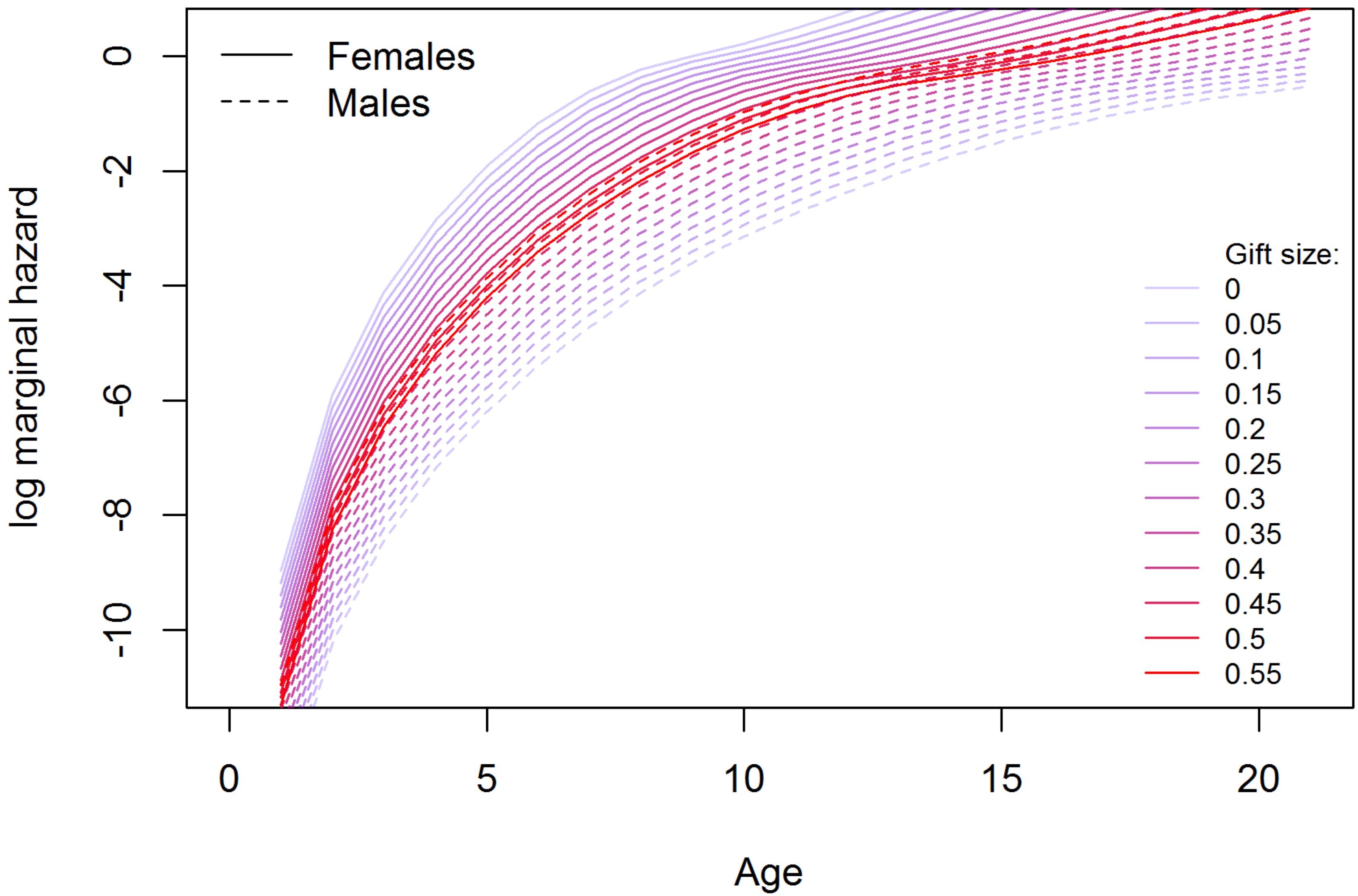
Effect of gift size on marginal hazard rate in males and females

My new extension to *parfm* package



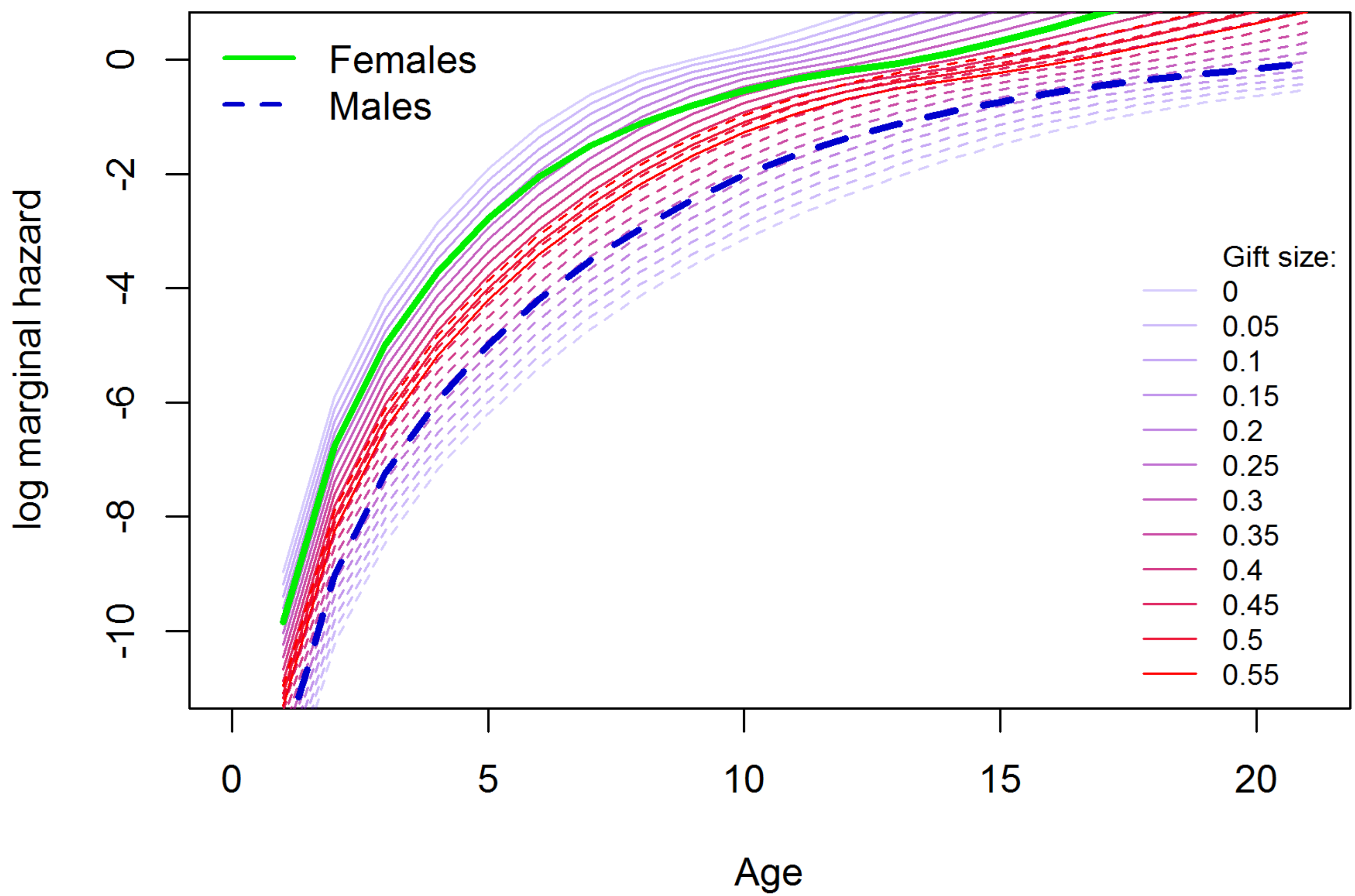
Effect of gift size on marginal hazard rate in males and females

My new extension to *parfm* package



Effect of gift size on marginal hazard rate in males and females

My new extension to *parfm* package



Summary

- The survival costs of reproduction are more pronounced for females than for males:
  - non-reproducing females live longer than non-reproducing males
  - Reproducing females live shorter than reproducing males
- Nuptial gifts increases substantially survival of the females at the costs of decreased survival of males.
- Bigger body mass is related to better survival in both males and females
- Size of the bean doesn't seem to influence survival of an adult

Summary

- The survival costs of reproduction are more pronounced for females than for males:
    - non-reproducing females live longer than non-reproducing males
    - Reproducing females live shorter than reproducing males
- Nuptial gifts increases substantially survival of the females at the costs of decreased survival of males.
- Bigger body mass is related to better survival in both males and females
- Size of the bean doesn't seem to influence survival of an adult

Is frailty term significant?
#1 Chisq=3.31, pval=0.0689
#2 Chisq=2.14, pval=0.1435