Bayesian Semiparametric Methods for Joint Modeling Longitudinal and Survival Data

Tim Hanson

Division of Biostatistics

University of Minnesota

and

Adam Branscum

University of Kentucky

Departments of Biostatistics, Statistics, and Epidemiology

October 11, 2006

Outline

- The joint modeling setting
- Why consider joint models?
- Data structure
- Background on modeling approaches
- Models for survival component
- Model for longitudinal component
- Bayesian semiparametric joint models
- Illustration
- Conclusions

Joint modeling setting

- Data collected from epidemiologic studies such as clinical trials or observational cohort studies often include information for an event time of interest (e.g. survival times) and repeated measurements of one or more longitudinal processes (e.g. exposure history) that might be associated with patient prognosis.
- Tsiatis and Davidian (2004) provide a detailed overview of statistical methodology for joint longitudinal/survival models focusing on the Cox (1972) model.
- Used to make inferences for two common study objectives:
  1. Trends in the time courses of longitudinal processes.
  2. Association between time-dependent variables and event prognosis.

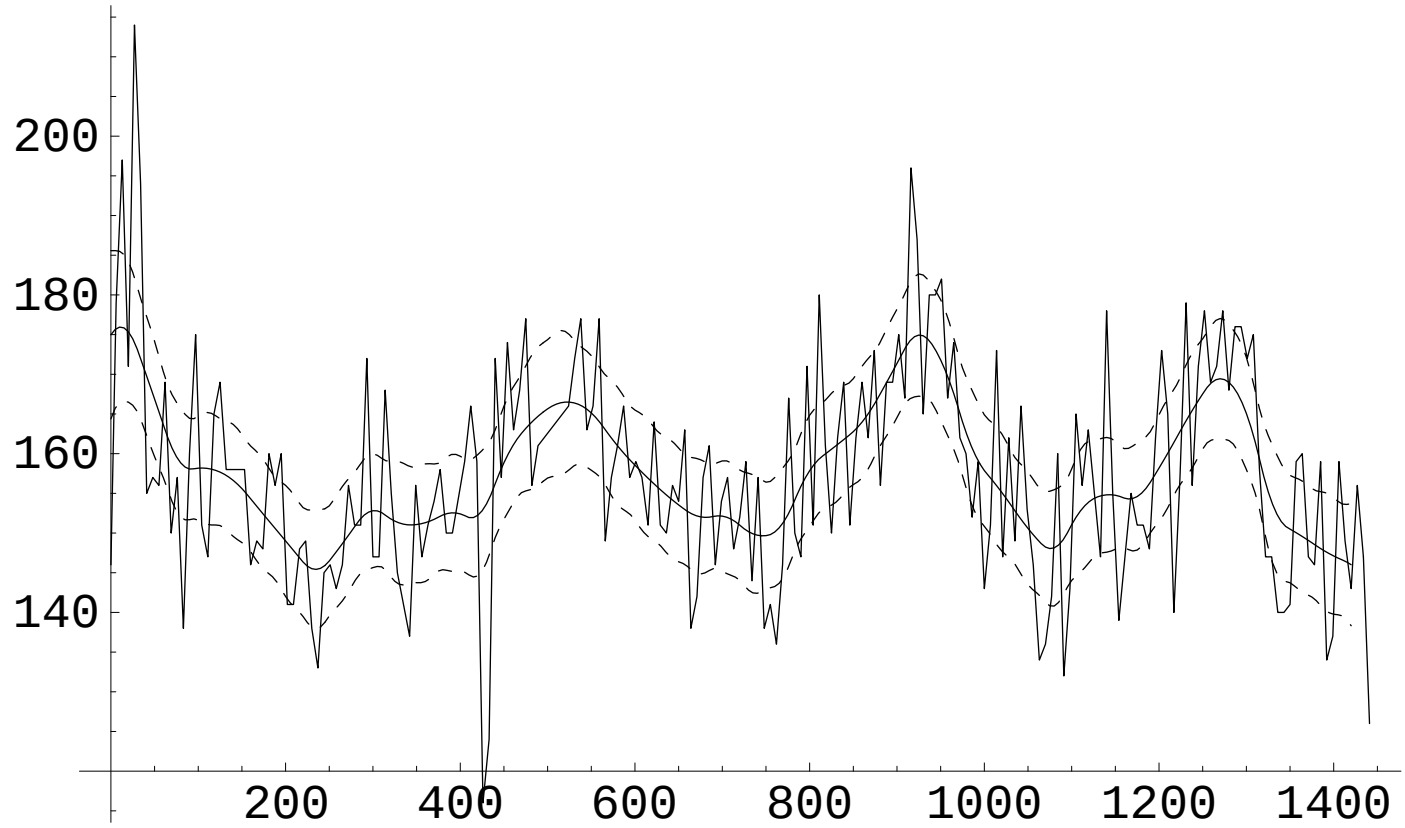


Figure 1: Weekly averaged systolic blood pressure over 1400 days for a hemodialysis patient until death. Quadratic B-spline fit ($d = 48$ functions on 47 equispaced knots). Q: anything here that signals death?

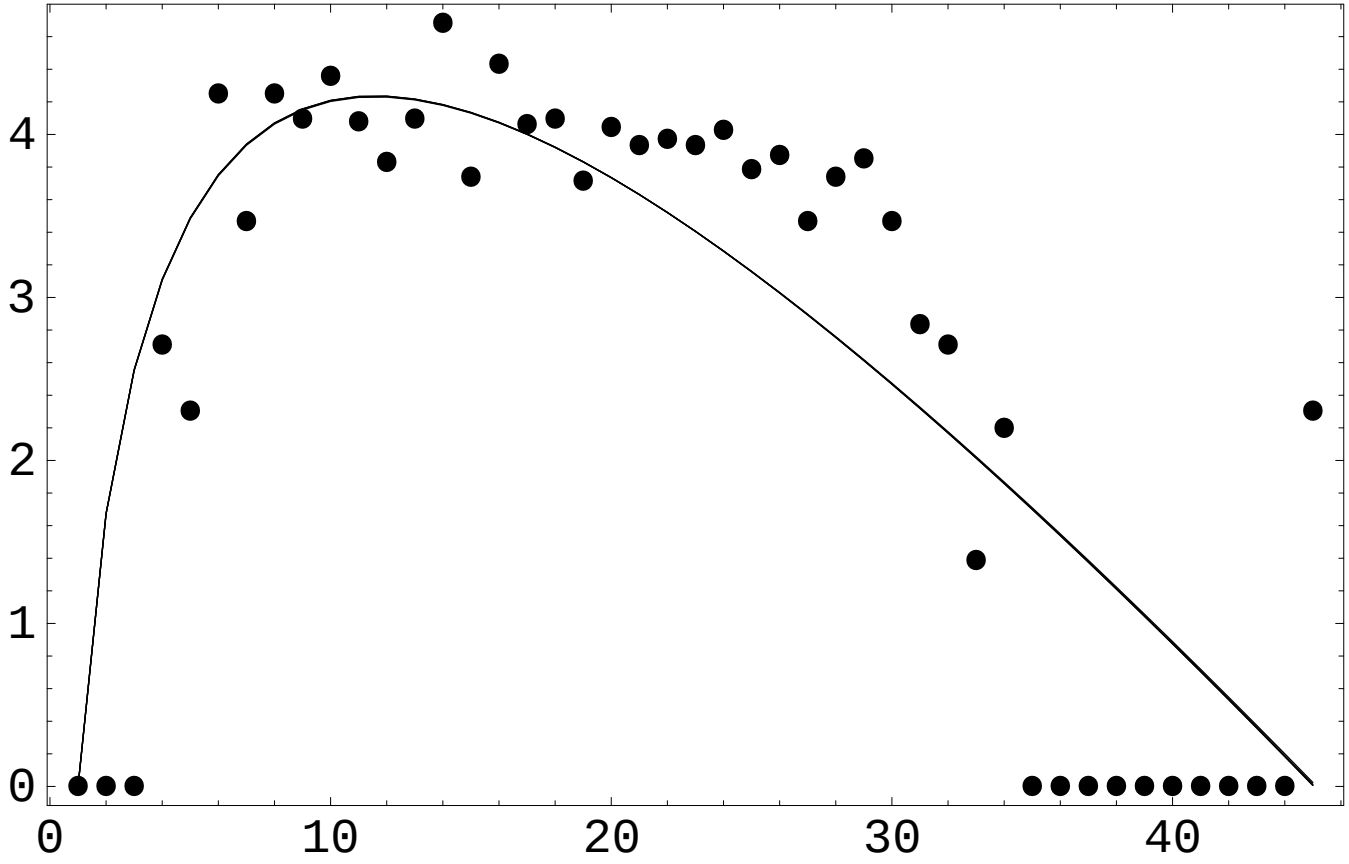


Figure 2: Log number of eggs laid on each day by medfly. Q: anything here that signals death?

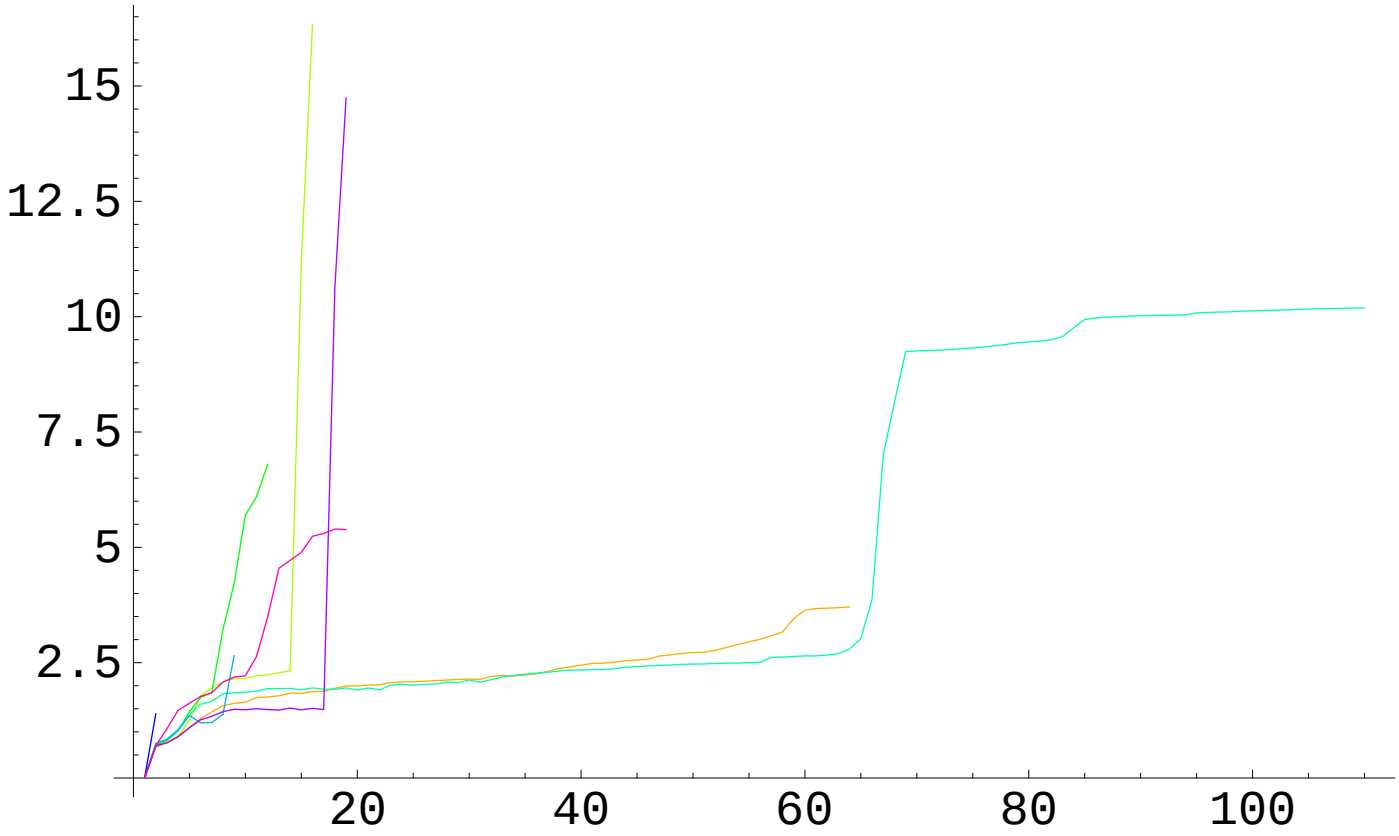


Figure 3: Resistance in $\mu\Omega$ of 8 ATC's before bond pad failure at temp=30^o C, humidity = 30% and NaCl = 20 μg cm^2 (time units = 12 seconds). Q: anything here that signals failure?

Why Consider Joint Modeling?

- Three alternatives to joint modeling include:
  1. Separate analyses for the survival and longitudinal processes.
  2. Survival analysis with time-dependent covariates (TDCs) using LVCF.
  3. Two-stage procedures.
- Two-stage procedures are conducted by imputing unobserved TD variable values by modeling the longitudinal process first and then using the imputed values in a TDC survival model.
- In our illustration, we compare joint analyses with these alternative approaches.

Data Structure

- The joint models we consider combine two linked submodels: a model for the longitudinal component and a model for the survival component.
- A longitudinal process, $x(\cdot)$, is measured with error so we observe $y(\cdot)$ at several time points where

$$y(t) = x(t) + \epsilon(t).$$

- A time to event, T_i, that is subject to right censoring is also observed for each sampled individual. The data for subject i, $i = 1, \ldots, n$, are denoted by

$$(T_i, \delta_i, \mathbf{y}_i, \mathbf{z}_i, \mathbf{t}_i).$$

- We assume independence of data across subjects i, and noninformative censoring and measurement schedule.

Examples:

- Log of daily number of eggs y_{ij} laid $j = 1, 2, 3, \ldots$ by Mediterranean fruit flies (medflies). Endpoint T_i is death of fly.
- Monthly $\mathbf{y}_{ij}$ =($debt, $assets) for a firm. Fixed covariates $\mathbf{z}_i$ might include type of firm. Endpoint T_i is filing for bankruptcy (death of firm?).
- (CD4,CD8) counts $\mathbf{y}_{ij}$ collected sporadically (two correlated processes) in HIV+ patients. Endpoint T_i is death of patient.
- Weekly systolic (pre) blood pressure of hemodialysis patient averaged over Monday, Wednesday, and Friday. Endpoint T_i is death of patient.
- Bond pad resistance y_{ij} in $\mu\Omega$ over time, collected sporadically, although "monitored" every 12 seconds. Fixed covariates $\mathbf{z}_i$ humidity and NaCl. Endpoint T_i is component failure.

Background: Longitudinal component

- Many previous approaches to joint modeling have used mixed effects models for the observed longitudinal process that are special cases of the following general structure:

$$\begin{aligned} y_i(t) &= x_i(t) + \epsilon_i(t) \\ x_i(t) &= \mathbf{f}(t)'\boldsymbol{\gamma} + \mathbf{g}(t)'\mathbf{b}_i + U_i(t) + \mathbf{z}_i(t)'\boldsymbol{\alpha} \\ \epsilon_i(t) &\overset{iid}{\sim} N(0, \sigma^2) \end{aligned}$$

  - $\mathbf{f}(t)$ and $\mathbf{g}(t)$ are vectors of known functions of time.
  - $U_i(t)$ is a mean-zero stochastic process, for example an Ornstein-Uhlenbeck process or IOU process.
  - The vector of random effects, $\mathbf{b}_i$, is typically modeled as multivariate normal $\perp$ of the $\epsilon_i(t)$'s and $\mathbf{b}_{-i}$.

Background: Survival component

- Let $x^H(t) = \{x(s) : s < t\}$ denote the history of the process $x(\cdot)$ up to time t.
- The majority of applications of joint modeling have used a Cox PH model for the survival component.
- The longitudinal and survival submodels are linked using one of the following methods:
  - The Cox model includes $x(\cdot)$ as a TD variable, e.g.

$$h(t|x^H(t)) = e^{x(t)\beta} h_0(t)$$

  - The longitudinal and survival components are associated using correlated stochastic processes: one SP in the longitudinal component and one in the survival component (e.g. Henderson et al. 2000).

Background: Survival component

- For the Cox model, the risk of failure at time t is assumed to depend on the current value of x, and not on its history.
- Although this assumption may be valid in some cases, a cumulative effect of biomarkers and/or exposure or treatment processes will be biologically appropriate in other cases.
- We consider a Bayesian treatment of a model developed by Cox and Oakes (1984) in which cumulative covariate effects up to time t influence survival prospects at time t, which will be biologically consistent in many settings that involve longitudinal exposures, disease biomarkers, or treatment covariate processes.

PH may be inappropriate

Say $x(t)$ indicates whether or not someone is currently smoking at time t, and this individual quits at time t_0:

$$x(t) = \left\{ \begin{array}{ll} 0 & t \geq t_0 \\ 1 & t < t_0 \end{array} \right\}.$$

then

$$h(t) = \left\{ \begin{array}{ll} h_0(t) & t \geq t_0 \\ e^{\beta} h_0(t) & t < t_0 \end{array} \right\}.$$

Instantaneous risk of dying immediately jumps back to that of nonsmoker at time t_0. Not realistic, especially if smoking several packs a day for 20 years.

Models for survival data

- We consider several survival models focusing on model selection via predictive performance. In particular:
  - Extending the AFT model of Cox and Oakes (1984) to a Bayesian joint specification.
  - Extending the proportional odds rate model of Sundaram (2006) to a Bayesian joint specification.
  - Comparison of competing joint models from a predictive standpoint.
- We also consider the Cox model in a joint analysis, models that treat the longitudinal process as a fixed TDC, and two-stage procedures.
- These methods are illustrated on a data set relating lifetimes of female fruit flies to daily egg production.

Models for survival data: AFT

- The AFT model with fixed-time covariates relates the event time rv T for a subject with covariate x to that for a baseline subject (i.e. a subject with $x = 0$) by

$$T = e^{-x\beta} T_0, \qquad T_0 \sim S_0$$

  where $S_0(t) = \Pr(T_0 > t)$ is a baseline survival function.

- Parametric approaches specify S_0 as, e.g., log-logistic.

- Hanson and Johnson (2002) generalized parametric approaches to allow for arbitrary S_0 that was modeled with a mixture of (finite) Polya trees (MFPT) prior.
  - Here, the prior is $S_0 \sim \int PT(c, G_\psi) p(d\psi, dc)$ where c is a weight parameter associated with the PT and G_ψ is a standard parametric AFT distribution (e.g. log-logistic).

Models for survival data: COAFT

- Cox and Oakes (1984) proposed a generalization of the AFT model that accommodates TDCs.
- The COAFT model assumes that an individual with covariate $x(\cdot)$ uses up their lifetime at a rate of $e^{x(t)\beta}$ relative to their counterfactual baseline rate:

$$\begin{aligned} T_0 &= \int_0^T e^{x(s)\beta} ds \\ T_0 &\sim S_0. \end{aligned}$$

- Tseng et al (2005) considered a semiparametric frequentist joint model that involved COAFT for the survival component and developed a Monte Carlo EM algorithm to fit the model, and bootstrap standard errors of regression coefficients were obtained.

Models for survival data: COAFT

- The hazard, survival, and density functions under the fixed-time covariate AFT model generalize to the following forms under the COAFT model:

$$\begin{array}{rcl} \underline{\text{AFT}} & & \underline{\text{COAFT}} \\ h(t|x) = e^{x\beta} h_0(e^{x\beta} t) & \rightarrow & h(t|x^H(t)) = e^{x(t)\beta} h_0(\bar{c}(t)t) \\ S(t|x) = S_0(e^{x\beta} t) & \rightarrow & S(t|x^H(t)) = S_0(\bar{c}(t)t) \\ f(t|x) = e^{x\beta} f_0(e^{x\beta} t) & \rightarrow & f(t|x^H(t)) = e^{x(t)\beta} f_0(\bar{c}(t)t) \end{array}$$

where

$$\bar{c}(t) = \frac{1}{t} \int_0^t e^{x(s)\beta} ds$$

is an "average acceleration factor" up to time t.

Models for survival data: PO

- With fixed-time covariates, the proportional odds (PO) model specifies

$$\frac{1 - S(t|x)}{S(t|x)} = e^{x\beta} \frac{1 - S_0(t)}{S_0(t)}.$$

- Sundaram (2006) extended this model to allow for TDCs yielding a POTDC model, which is defined by

$$\frac{d}{dt}\left[\frac{1 - S(t|x(\cdot))}{S(t|x(\cdot))}\right] = e^{x(t)\beta} \frac{d}{dt}\left[\frac{1 - S_0(t)}{S_0(t)}\right].$$

- Model is in terms of odds rate, β difficult to interpret. Need to talk about relative rate at which odds of dying before t are increasing.

Comments

- TDC versions of PH, AFT, and PO reduce to standard PH, AFT, and PO models when $x(t) \equiv x_0$, a constant.
- When modeling S_0, possible to obtain Cox-Snell residuals and make plots to assess model fit.
- Can add multiple TDCs and fixed covariates as well.
- Omitted Aalen (1980) semiparametric model:

$$h(t|x(\cdot)) = h_0(t) + x(t)\beta$$

  hard to fit in joint modeling context, very little done here in either Bayesian or frequentist realms.

Models for survival data: TDC

In the illustration that follows, the COAFT, PO, and PH models are considered with $x(\cdot)$

- treated as a fixed time dependent covariate using LVCF,
- modeled using an expansion of simple basis functions, and
- imputed from fitting smooth expansion.

Model for longitudinal data

- Generally decided on a case-by-case basis.
- To be consistent with a previous joint analysis of the medfly data (Tseng et al, 2005), we consider the following structure.
- Where $\mathbf{y}_i = (y_{i1}, \ldots, y_{in_i})'$ are the n_i longitudinal measurements of subject i at times $\mathbf{t}_i = (t_{i1}, \ldots, t_{in_i})'$, the hierarchical model specifies that trajectories satisfy

$$y_{ij}|\mathbf{b}_i, \sigma^2 \stackrel{\perp}{\sim} N\left(b_{i1}g_1(t_{ij}) + b_{i2}g_2(t_{ij}) + \cdots + b_{id}g_d(t_{ij}), \sigma^2\right)$$

- We assume the individual trajectories are iid from a family of such curves,

$$\mathbf{b}_i|\boldsymbol{\mu}, \boldsymbol{\Sigma} \stackrel{iid}{\sim} N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}).$$

The assumptions on the individual curves

$$\mathbf{b}_i|\boldsymbol{\mu}, \boldsymbol{\Sigma} \overset{iid}{\sim} N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}).$$

can be relaxed in two directions.

The trajectories can be functions of fixed covariates (i.e. a multilevel model), e.g.

$$\mathbf{b}_i|\mathbf{B}, \boldsymbol{\Sigma} \overset{iid}{\sim} N_d(\mathbf{B}'\mathbf{z}_i, \boldsymbol{\Sigma}).$$

The normality assumption can be relaxed, e.g.

$$\mathbf{b}_i|G \overset{iid}{\sim} G, \qquad G|\boldsymbol{\mu}, \boldsymbol{\Sigma} \sim DP(cN_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})).$$

An example of the former is Brown, Ibrahim, and DeGruttola (2005), the latter, Brown and Ibrahim (2003). The latter generally has little effect on LMPL.

Model fitting

- Let $x_i(t|\mathbf{b}_i) = b_{i1}g_1(t) + \cdots + b_{id}g_d(t)$ denote the latent mean trajectory of subject i's longitudinal measurements.
- For joint models, the survival component is specified conditional on longitudinal processes $\{x_i(\cdot|\mathbf{b}_i)\}_{i=1}^n$
- We allow S_0 to be arbitrary and place an MFPT prior (Hanson, 2006) on it with
  - log-logistic centering family, i.e. $E(S_0(t)) = (1 + t^{1/\tau}e^{-\alpha/\tau})^{-1}$,
  - collection of branch probabilities $\mathcal{X}_M$ & weight parameter c.
- Let $\boldsymbol{\theta} = (\alpha, \tau, \mathcal{X}_M, c)$.
- A model $[T_i|\boldsymbol{\theta}, \beta, x_i(\cdot|\mathbf{b}_i)]$ is specified as COAFT, PO, or PH, and longitudinal measurements $\mathbf{y}_i$ are further modeled according to $\mathbf{y}_i|\mathbf{b}_i, \sigma \sim N_{n_i}(\mathbf{X}_i\mathbf{b}_i, \mathbf{I}_{n_i}\sigma^2)$.

Model fitting

- Independent priors:
  - $p(\mu, \mathbf{\Sigma}, \beta, \alpha, \tau) \propto |\mathbf{\Sigma}|^{-(d+1)/2}$
  - $p(\sigma^{-2}) \propto 1/\sigma^{-2}$
  - $c \sim \Gamma(a_c, b_c)$
  - $(X_{j,2k-1}, X_{j,2k}) \sim \text{Dirichlet}(cj^2, cj^2)$
- The posterior based on the survival portion, the longitudinal portion, and the prior is then

$$p(\beta, \boldsymbol{\theta}, \boldsymbol{\mu}, \mathbf{\Sigma}, \sigma | \mathbf{T}, \mathbf{y}_{1:n}) =$$
$$\left[\prod_{i=1}^{n} f(T_i | x_i(\cdot|\mathbf{b}_i), \boldsymbol{\theta}, \beta)^{\delta_i} S(T_i | x_i(\cdot|\mathbf{b}_i), \boldsymbol{\theta}, \beta)^{1-\delta_i}\right] \times$$
$$\left[\prod_{i=1}^{n} p(\mathbf{y}_i | \mathbf{b}_i, \sigma) p(\mathbf{b}_i | \boldsymbol{\mu}, \mathbf{\Sigma})\right] p(\beta, \boldsymbol{\theta}, \boldsymbol{\mu}, \mathbf{\Sigma}, \sigma)$$

Model fitting

- The full conditional distributions for $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, and σ^{-2} are:

$$\begin{aligned}
\boldsymbol{\Sigma}^{-1}|\mathbf{b}_{1:n}, \boldsymbol{\mu} &\sim \text{Wishart}\left(n, \left[\sum_{i=1}^{n}(\mathbf{b}_i - \boldsymbol{\mu})(\mathbf{b}_i - \boldsymbol{\mu})'\right]^{-1}\right) \\
\boldsymbol{\mu}|\mathbf{b}_{1:n}, \boldsymbol{\Sigma} &\sim N_d\left(\bar{\mathbf{b}}_\bullet, \boldsymbol{\Sigma}/n\right) \\
\sigma^{-2}|\mathbf{b}_{1:n} &\sim \Gamma\left(0.5\sum_{i=1}^{n} n_i, 0.5\sum_{i,j}(y_{ij} - x_i(t_{ij}|\mathbf{b}_i))^2\right)
\end{aligned}$$

- Metropolis-Hastings steps were used to sample the full conditionals for the $\mathbf{b}_i$'s (updated w/ proposal based on longitudinal model or random-walk M-H; both gave same inferences), $\mathcal{X}_M$ (w/ beta proposals), c (w/ truncated normal proposal), (α, β, τ) (w/ random walk M-H).

Comment:

The linear model is one approach to modeling $x_i(t|\mathbf{b}_i)$. Alternatives include:

- Nonlinear regression models, e.g.

$$y_{ij} = \frac{b_{i1}}{1 + b_{i2} b_{i3}^{t_{ij}}} + \epsilon_{ij} = x_i(t_{ij}|\mathbf{b}_i) + \epsilon_{ij}.$$

- Changepoint models, e.g.

$$y_{ij} = b_{i1} + b_{i3}(t_{ij} - b_{i2})^+ + \epsilon_{ij} = x_i(t_{ij}|\mathbf{b}_i) + \epsilon_{ij}.$$

  Here, $(t)^+$ is t if $t \geq 0$ and zero otherwise.

However, the linear model includes unpenalized spline models, wavelet expansions, etc., for fixed d. Penalized spline models can similarly be fit with special structure on $\mathbf{\Sigma}$.

Model choice

- We compare parametric and semiparametric joint analyses to survival analysis with fixed TDC, and to two-stage models.
- The model selection criterion we use, the log-pseudo marginal likelihood (LMPL), compares predictive ability of failure time among competing models:

$$\text{LPML} = \sum_{i=1}^{n} \log(\text{CPO}_i)$$

  where $\text{CPO}_i = p(T_i|\mathbf{T}_{-i}, \mathbf{y}_{1:n})$

- Although not used in our example, one can also consider an LPML measure that focuses on both prediction of survival and the longitudinal trajectory

Model choice

- We can calculate CPO_i from MCMC output because

$$
\begin{aligned}
& E\left[\frac{1}{p(T_i|\mathbf{b}_i,\beta,\boldsymbol{\theta})}\right] \\
&= \int \frac{p(\mathbf{b}_{1:n},\beta,\boldsymbol{\theta}|\mathbf{T},\mathbf{y}_{1:n})d\mathbf{b}_{1:n}d\beta d\boldsymbol{\theta}}{p(T_i|\mathbf{b}_i,\beta,\boldsymbol{\theta})} \\
&= \int \frac{\left[\prod_{j\neq i} p(T_j|\mathbf{b}_j,\beta,\boldsymbol{\theta})\right]p(\mathbf{y}_{1:n}|\mathbf{b}_{1:n})p(\mathbf{b}_{1:n})p(\beta,\boldsymbol{\theta})d\mathbf{b}_{1:n}d\beta d\boldsymbol{\theta}}{p(\mathbf{T},\mathbf{y}_{1:n})} \\
&= 1/\text{CPO}_i
\end{aligned}
$$

where expectation is taken wrt $[\mathbf{b}_{1:n},\beta,\boldsymbol{\theta}|\mathbf{T},\mathbf{y}_{1:n}]$ and

$$
\begin{aligned}
p(\mathbf{y}_{1:n}|\mathbf{b}_{1:n}) &= \int \prod_{i=1}^{n}\prod_{j=1}^{n_i} p(y_{ij}|\mathbf{b}_i,\sigma)P(d\sigma) \\
p(\mathbf{b}_{1:n}) &= \int \prod_{i=1}^{n} p(\mathbf{b}_i|\boldsymbol{\mu},\boldsymbol{\Sigma})P(d\boldsymbol{\mu},d\boldsymbol{\Sigma})
\end{aligned}
$$

Illustration: Medfly Data

- The data used for illustration came from a study where the reproductive patterns of 1000 female Mediterranean fruit flies (referred to as medflies) were obtained by recording the number of eggs produced each day throughout their lifespans.
- A scientific goal was to examine the association between egg production patterns and lifetime.
- A frequentist approach was used to analyze these data by Tseng et al (2005), and like these authors we excluded from our analyses flies whose lifetime egg production was < 1145.
- This gave a sample size of 251 flies with lifespans ranging from 22 to 99 days, and no censored observations.

Illustration: Medfly Data

- For joint models we used the longitudinal structure for egg production that was used by Tseng et al. The response was $\ln(y_i(t)+1)$ with

$$x_i(t|\mathbf{b}_i) = b_{1i}\log(t) + b_{2i}(t-1)$$

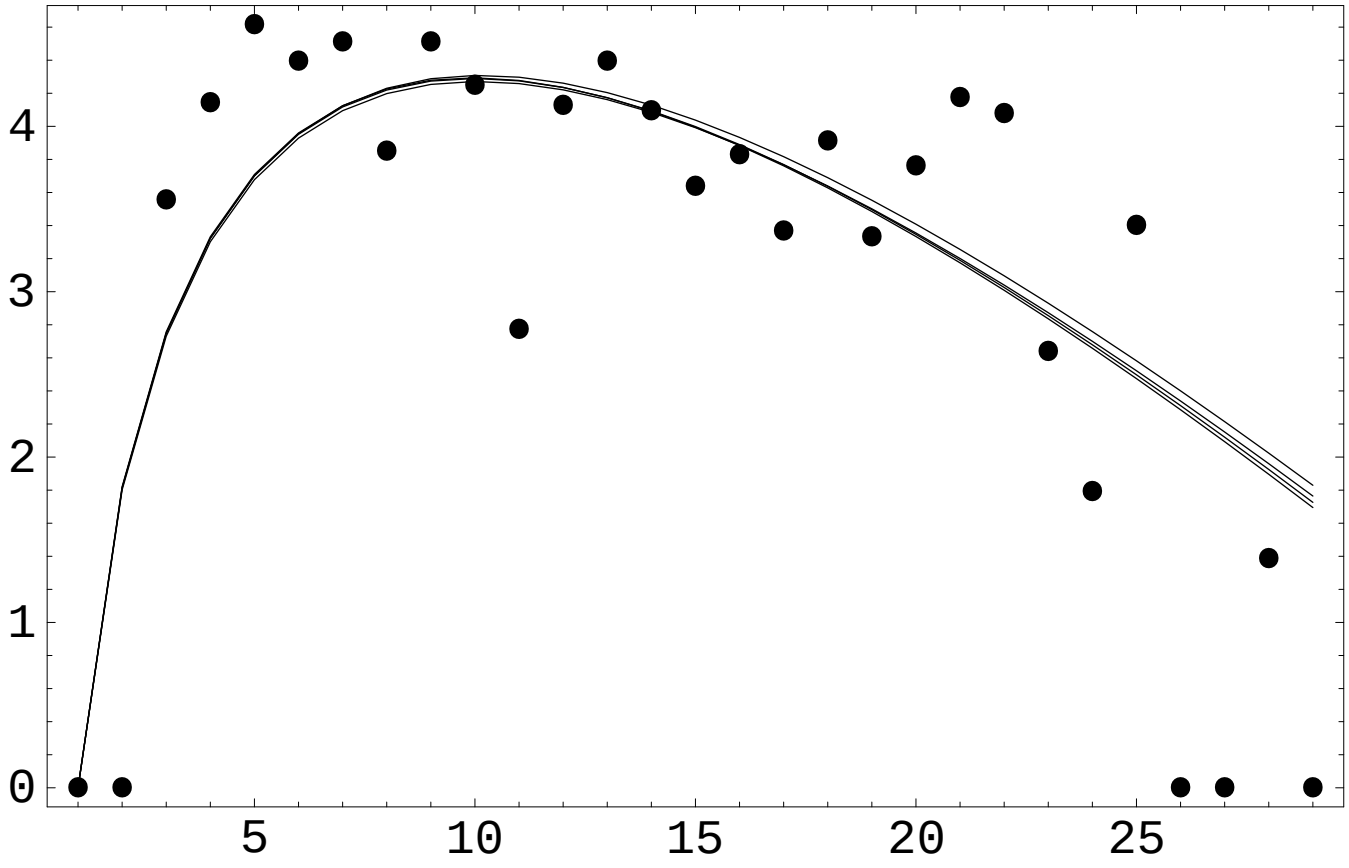
- For analyses that treated egg production as a fixed TDC, we used a piecewise constant function.
- For two-stage imputation approaches, we modeled the longitudinal data separately to obtain an estimated smooth temporal trend, and then fit survival models treating the estimated $x(\cdot)$ as a known TDC.

- LPML statistics comparing modeling approaches:

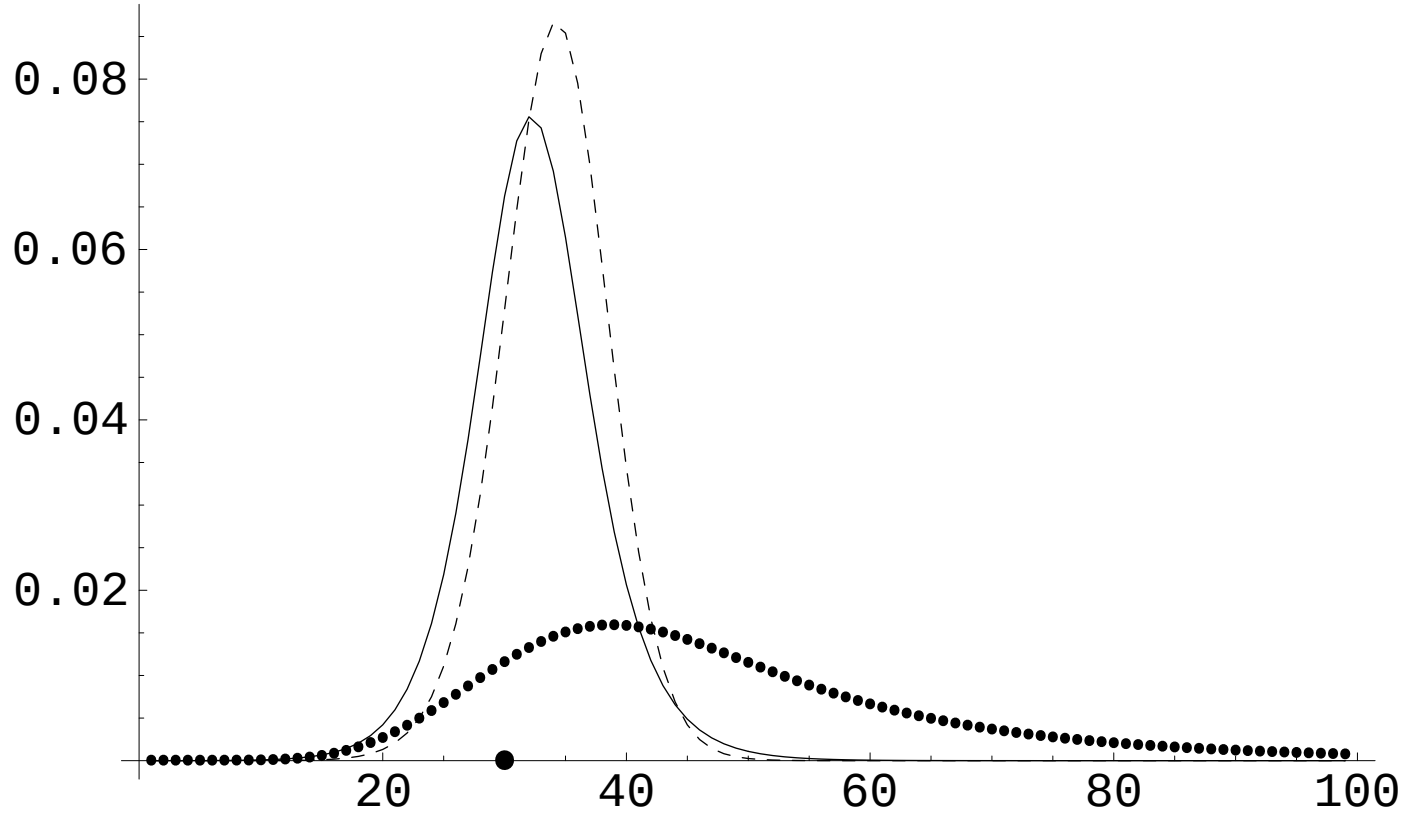
	PO	PH	CO
raw + parametric	−867	−870	−937
raw + MPT	−865	−866	−938
imputed + parametric	−947	−959	−973
modeled + parametric	−947	−959	−973
modeled + MPT	−945	−956	−973

- Summary based on LPML criterion:
  - Predictively, PO and PH models preferred over CO.
  - Survival with TDC using LVCF predictively better than joint analysis.
  - MPT improves predictive performance only slightly compared to parametric model.

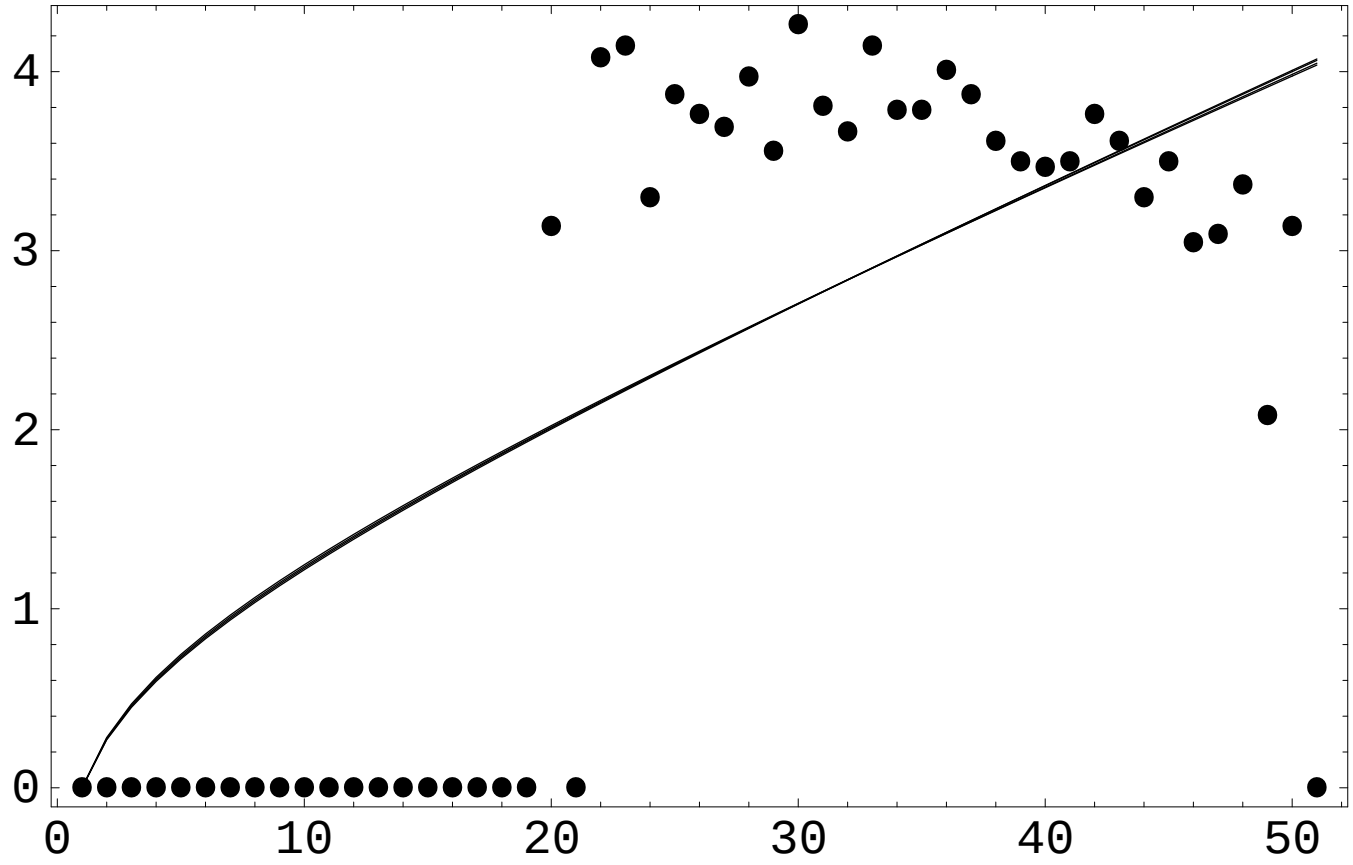
- **Fly 1**: fitted trajectory for a "typical" medfly. Similar shapes for PO, PH, CO, and longitudinal only analyses.



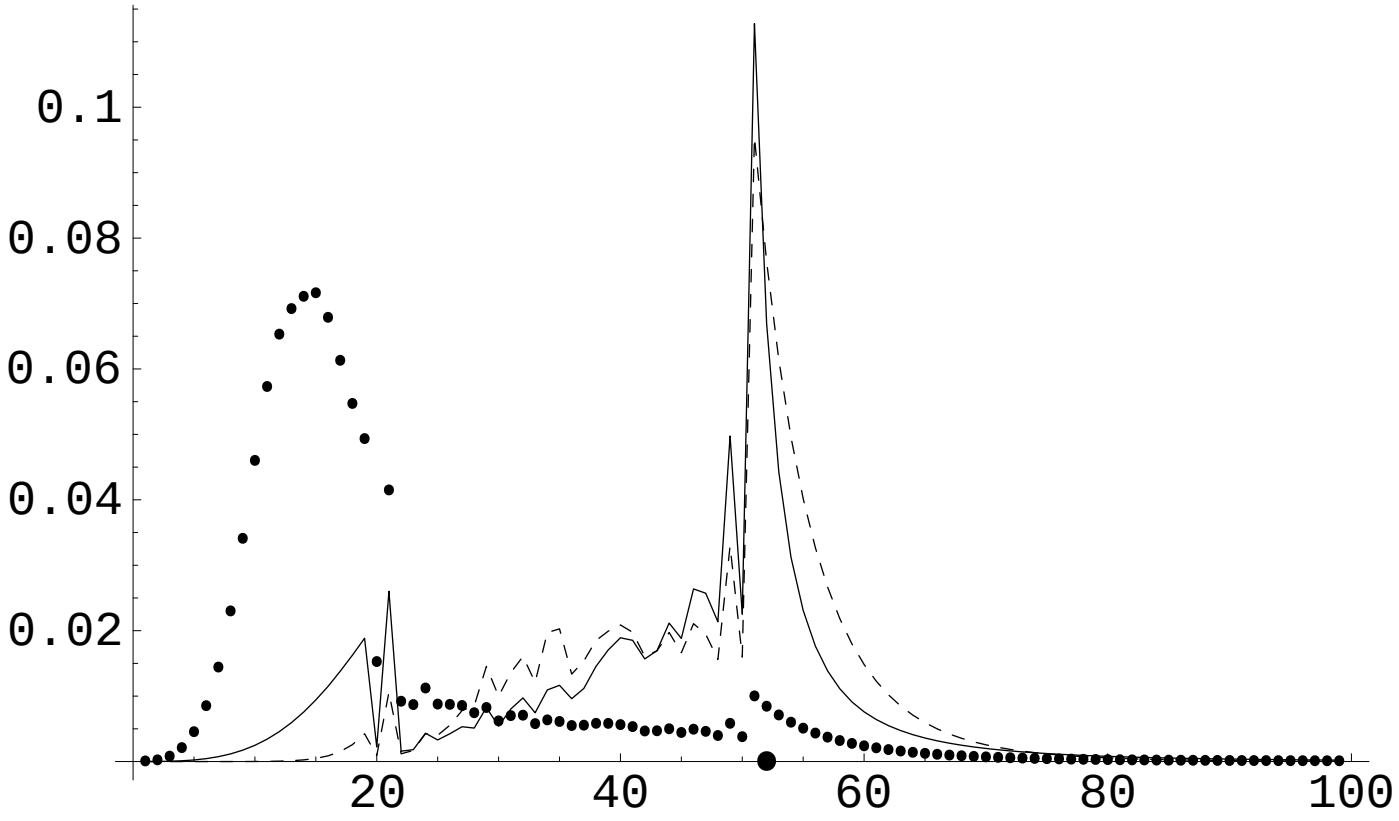
- **Fly 1**: predictive survival density, parametric models. Solid is PO, dashed is PH, and dotted is CO.



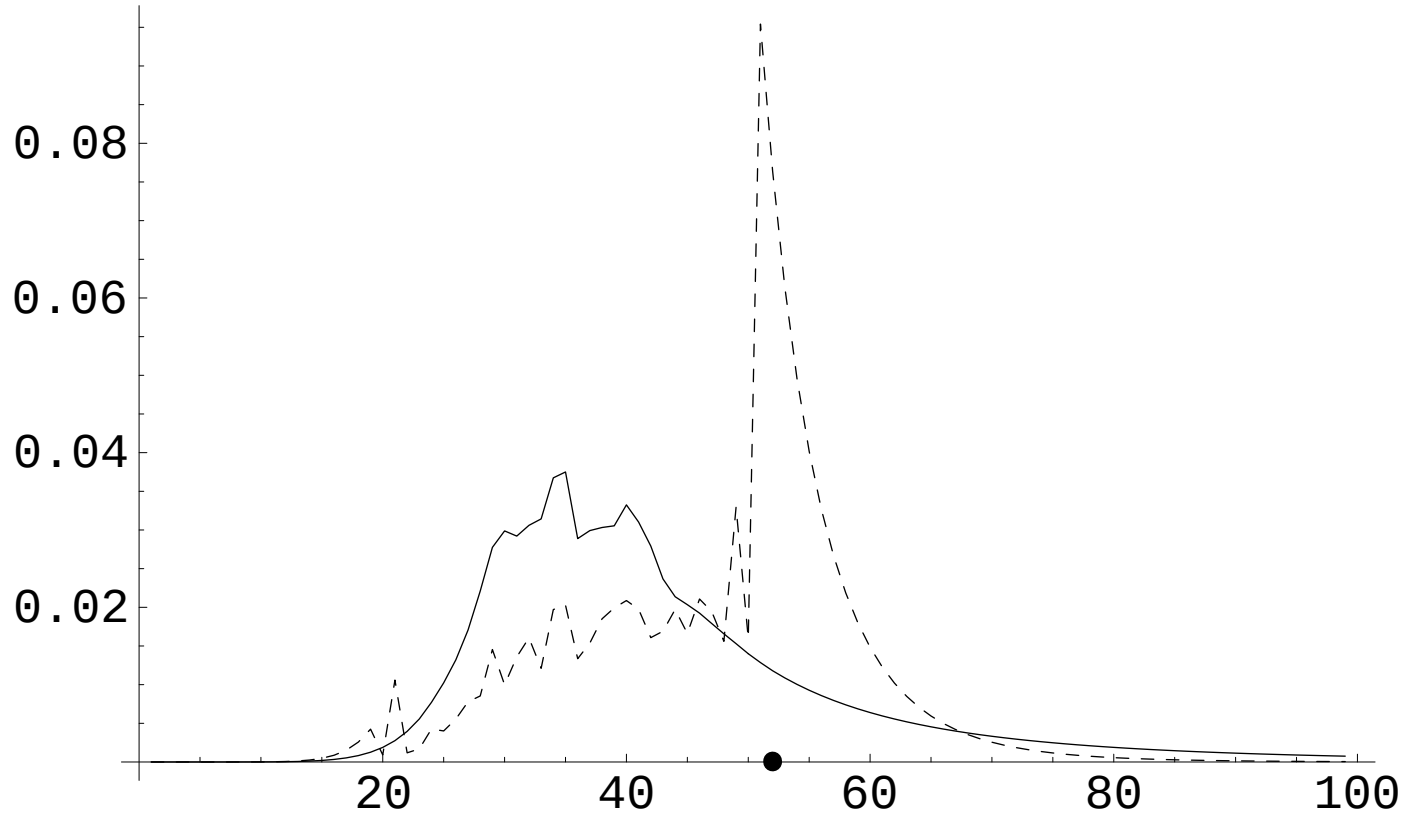
- **Fly 2**: fitted trajectory for another medfly using PO, PH, CO, and longitudinal only analyses.



- **Fly 2**: Predictive survival densities, semiparametric PO (solid), PH (dashed) and CO (dotted) analyses using raw trajectories.



- **Fly 2**: Predictive survival densities, semiparametric PH analyses comparing raw trajectories (dashed line) to joint analysis (solid line).



Posterior inference for β

	PO	PH	CO
raw + parametric	-0.75	-0.65	-0.36 (-0.44,-0.27)
raw + MPT	-0.74	-0.64	-0.37 (-0.45,-0.29)
imputed + parametric	-0.74	-0.37	0.16 (-0.01,0.30)
modeled + parametric	-0.78	-0.39	0.19 (0.01,0.33)
modeled + MPT	-0.79	-0.40	0.19 (0.01,0.32)

- $\Pr(\beta < 0 | \mathbf{T}, \mathbf{y}_{1:n}) = 1$ for PO and PH models $\Rightarrow$ survival prospects are better for the most fertile flies.
- For CO, inferences are different for joint models than for models based on raw trajectories.
  This may be due in part to relatively poor fits of temporal egg production for some medflies (e.g. Fly 2).

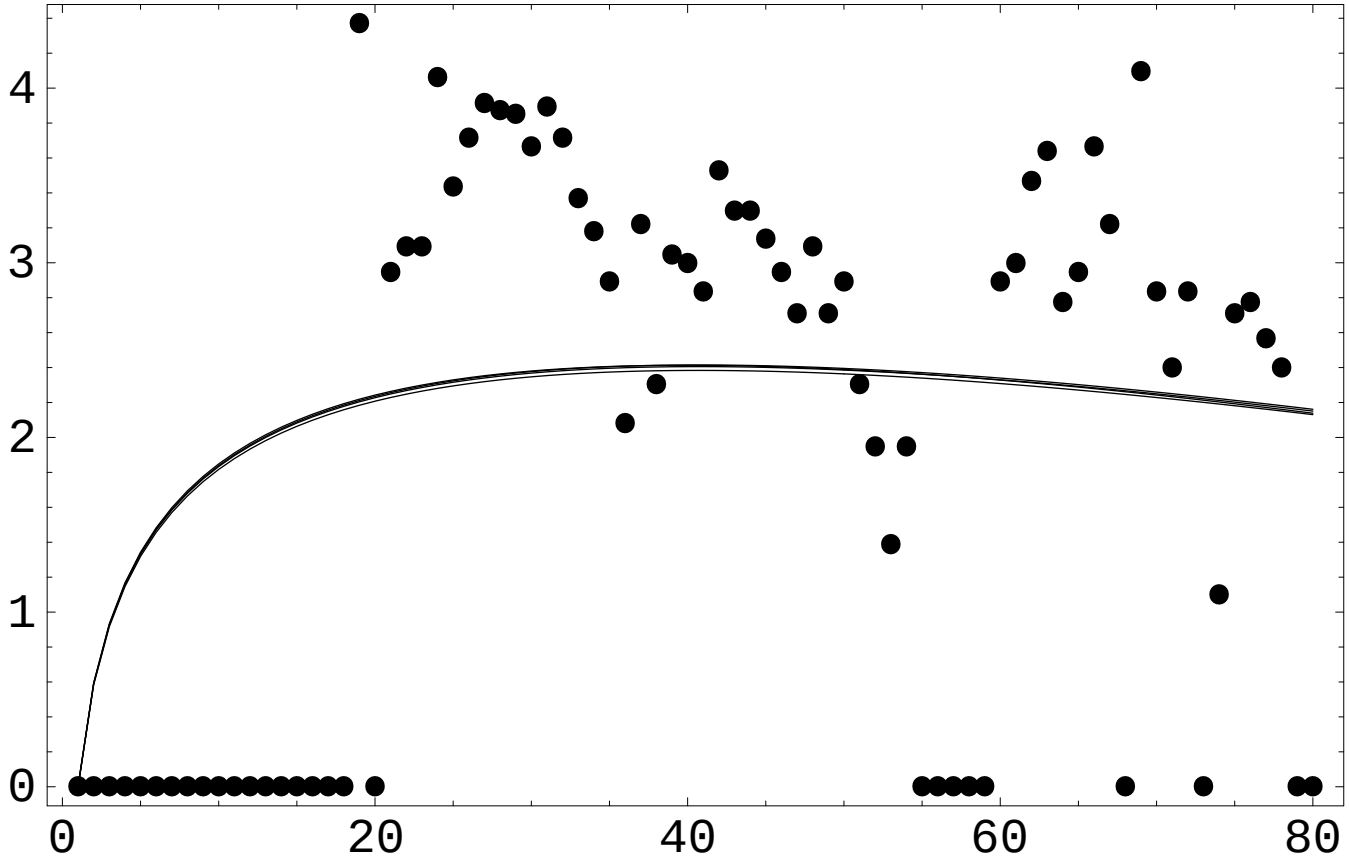
Comments:

- Whole point of Tseng et al. (2005) was that PH was inappropriate for the medfly data. They propose joint model combining nonparametric COAFT and longitudinal model as alternative. Made a big deal about choice of survival model and choice of basis.
- For medfly data, COAFT by far the worst model.
- The basis functions they picked yield worse inferences than just leaving the TDCs alone and using LVCF.
- LVCF may best when there's lots of data on each trajectory. Different than HIV/AIDS clinical trials where CD4 counts often sparsely collected and wildly variable over time.
- Modeling S_0 nonparametrically doesn't add anything for medfly data.

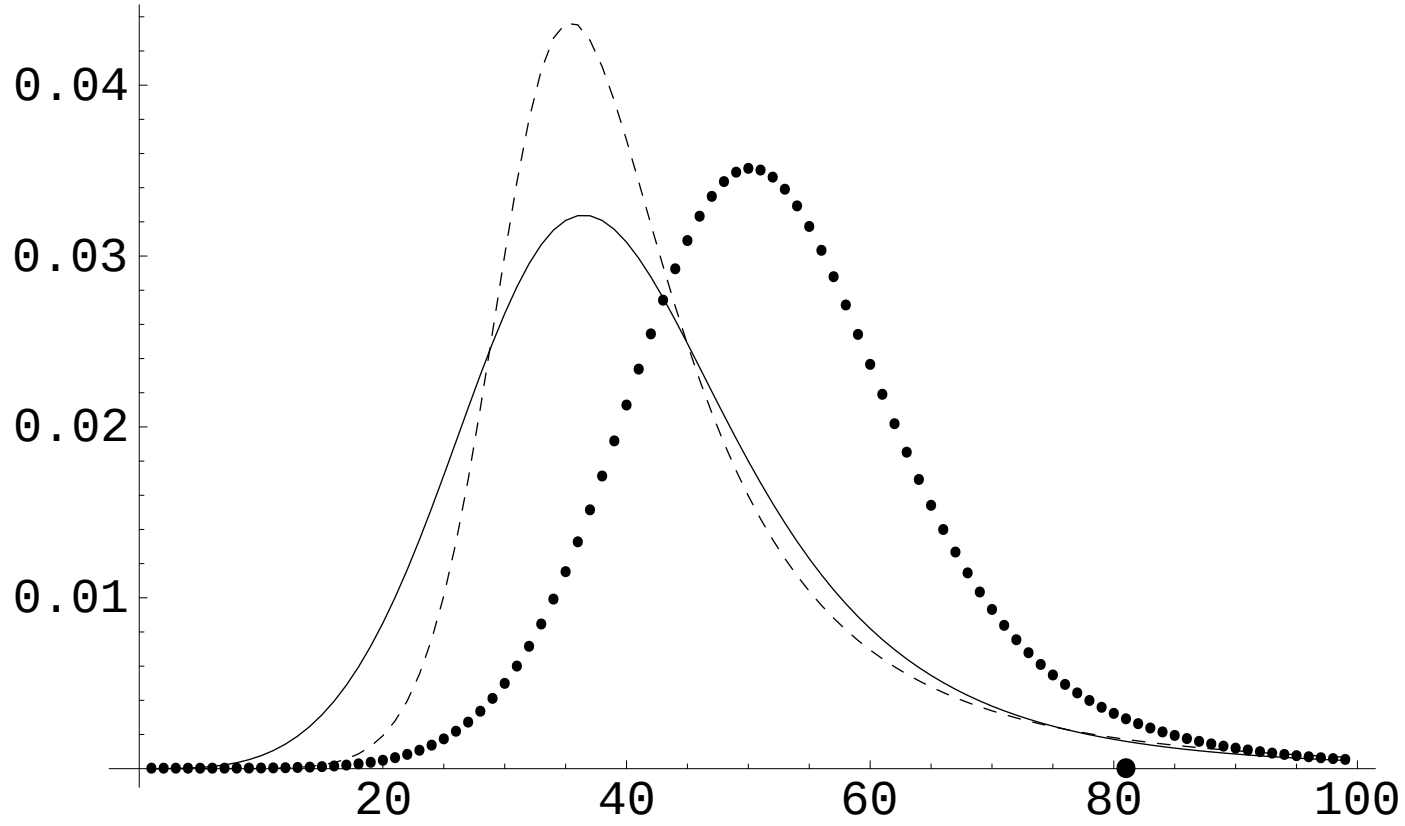
In other words...

- The survival model affects inference. You might be able to do a lot better with an alternative model. Might be a good idea to actually fit some alternatives.
- The longitudinal model affects inference. LVCF biases $\hat{\beta}$ in the PH model, but if longitudinal model is wrong, you could make things *worse*.
- Although nonparametric modeling adds flexibility, it may not help prediction.

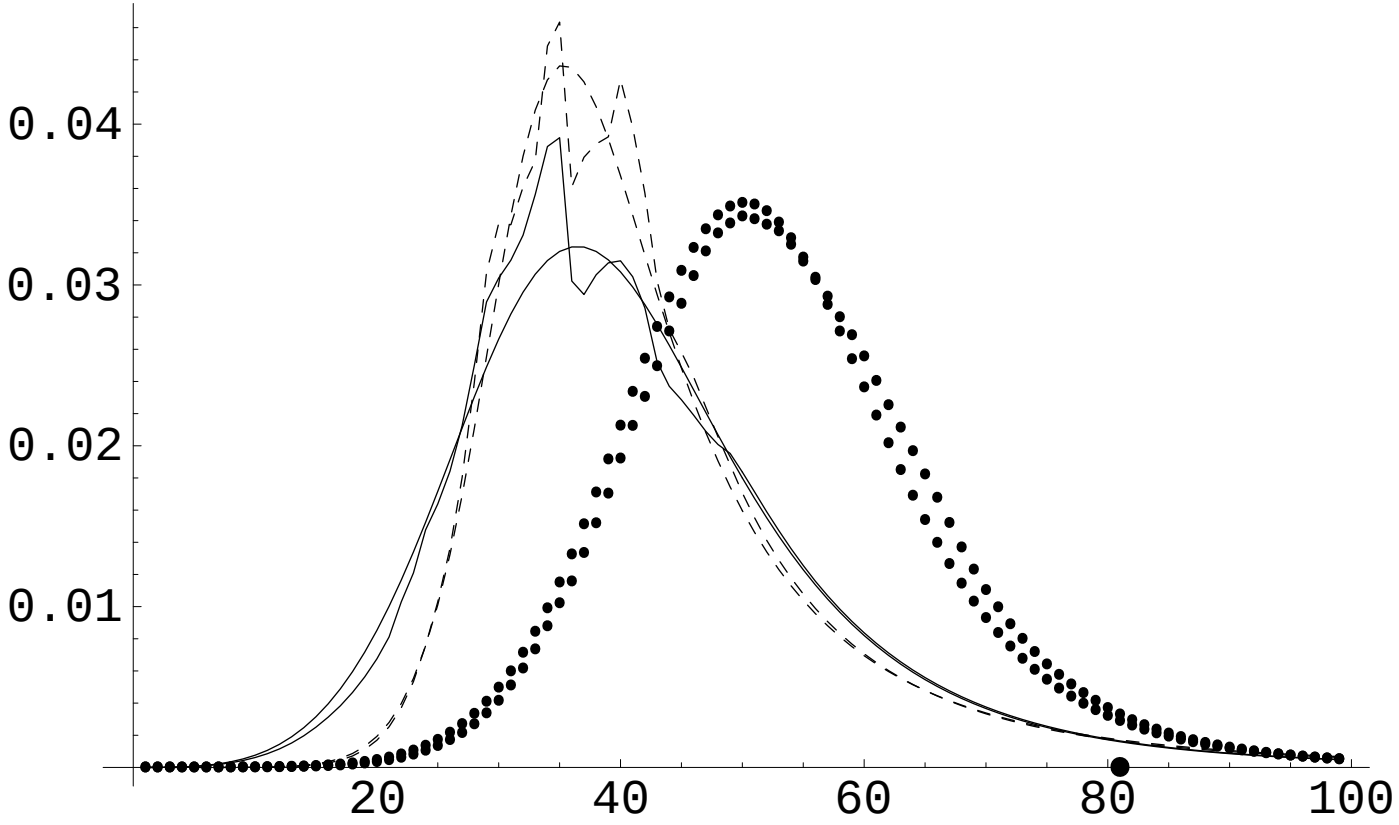
- **One more, fly 3**: raw data and fitted trajectories.



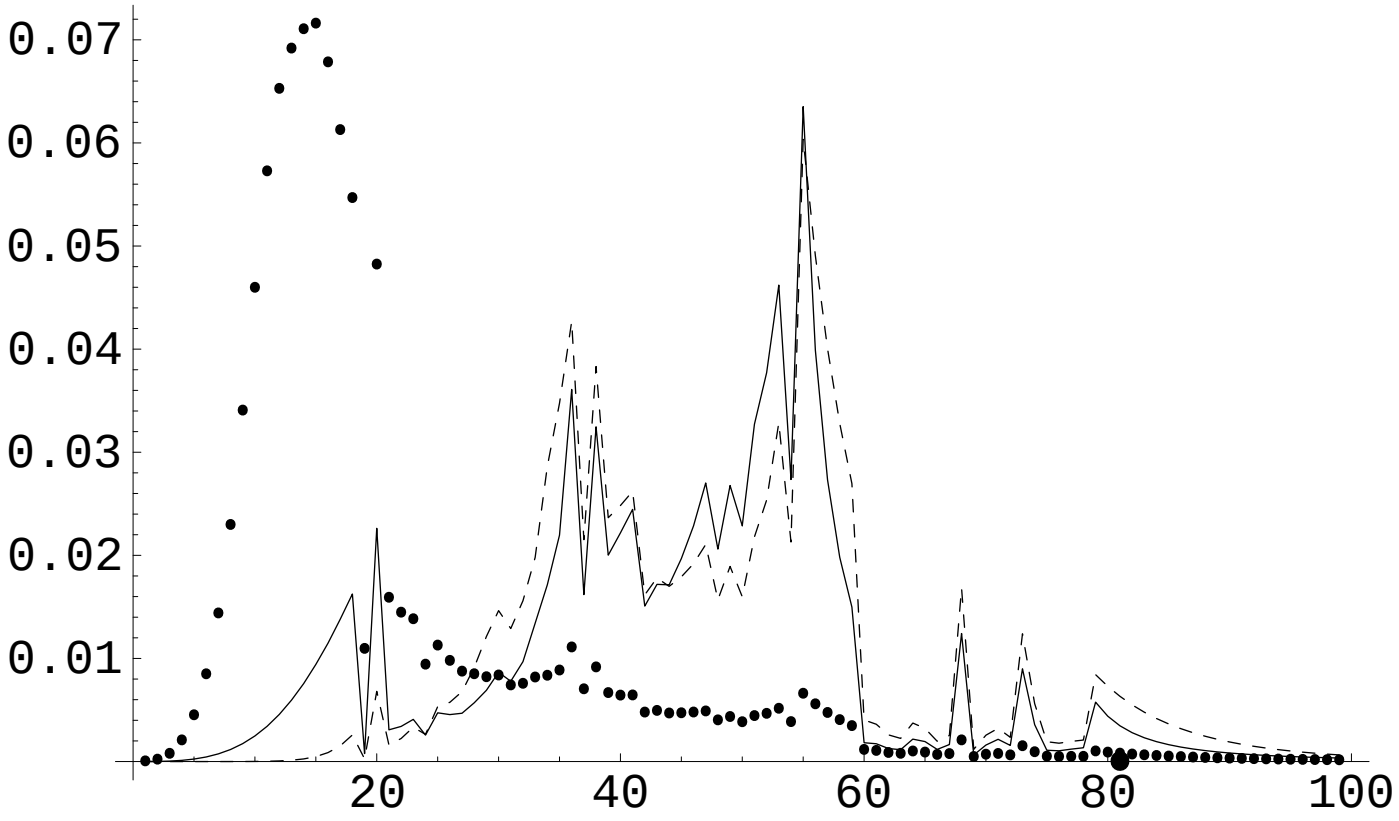
- **Fly 3**: Modeled trajectories, parametric model. Solid is PO, dashed is PH, and dotted is CO. (CO has largest CPO too).



- **Fly 3**: Modeled trajectories, parametric and MPT models. Solid is PO, dashed is PH, and dotted is CO. Nonparametric doesn't help much.



- **Fly 3**: Raw trajectories (LVCF), MPT models. Solid is PO, dashed is PH, and dotted is CO. (Now PH has largest CPO).



Discussion

- Bayesian semiparametric mixtures of Polya trees provide for flexible inference in joint regression modeling settings, however not needed here. Same baseline for S_0 in all models.
- Have found COAFT to be superior in data on time to bankruptcy of firms, PH to be superior in time to death of hemodialysis patients, PH superior in Stanford Heart Transplant data, etc.
- The survival component of joint models can be specified as COAFT, PH, or PO.
- Model selection can be carried out using LPML statistics and corresponding pseudo Bayes factors.

- Future work involves
  - More complex structures for $y(t)$; e.g.

    $$\mathbf{y}_i \sim N_{n_i}(\mathbf{X}_i\mathbf{b}_i, \mathbf{I}_{n_i}\sigma^2 + \mathbf{K}(\boldsymbol{\kappa})),$$

    where $\mathbf{K}(\boldsymbol{\kappa})$ obtained from mean-zero Gaussian process or penalized spline on top of trend $\mathbf{X}_i\mathbf{b}_i$.
  - AH can be fit but not pleasant to do. More work to be done.
  - New model, proportional mean life:

    $$E(T - t|T > t) = e^{x(t)\beta}E(T_0 - t|T_0 > t).$$

    Nice interpretation for β. TDC version hasn't been done.
- Was originally going to also talk about joint modeling of electrical component lifetimes as well. Main conclusion: raw trajectories work about as well as Gaussian process and IOU process on top of trend. Harder to predict into future though.

References

- Brown E.R. and Ibrahim J.G. (2003). A Bayesian semiparametric joint hierarchical model for longitudinal and survival data. *Biometrics*, 59, 221-228.
- Brown, E.R., Ibrahim, J.G., and Degruttola, V.A. (2005). Flexible B-spline model for multiple longitudinal biomarkers and survival. *Biometrics*, 61, 64-73.
- Cox, D.R. and Oakes, D. (1984). *Analysis of Survival Data.* Chapman and Hall, London.
- Hanson, T.E. (2006). Inference for mixtures of finite Polya tree models. *Journal of the American Statistical Association*, in press.
- Hanson, T. and Johnson, W.O. (2002). Modeling regression error with a mixture of Polya trees. *Journal of the American Statistical Association*, 97, 1020-1033.
- Henderson, R., Diggle, P.J., Dobson, A. (2000). Joint modelling of longitudinal measurements and event time data. *Biostatistics* 4, 465-480.
- Sundaram, S. (2006). Semiparametric inference in proportional odds model with time-dependent covariates. *Journal of Statistical Planning and Inference*, 136, 320-334.
- Tseng, Y.-K., Hsieh, F. and Wang, J.-L. (2005). Joint modelling of accelerated failure time and longitudinal data. *Biometrika*, 92, 587-603
- Tsiatis, A.A. and Davidian, M. (2004). Joint modeling of longitudinal and time-to-event data: an overview. *Statistica Sinica*, 14, 809-834.