General Dependence Diagnostics for Multivariate Failure Time Data

Dave Glidden

Univ. of California, San Francisco

October 5, 2005

UCLA

http://www.epibiostat.ucsf.edu/dave/talks.html

Motivation

- Interest: Familial aggregation of disease
- Outcome: age at onset
  unaffected subjects censored
  acknowledges importance of early onset
- Aggregation:
  existence, strength, implications?

Multivariate Failure Time

- Example: Ages of appendectomy
  Duffy, Australian Twin Study (1990)
- Data sampled in clusters (families)
  $i = 1, \ldots, n$
- Clusters contain failures $(T_1, \ldots, T_K)$
- How to describe, model dependence?

Frailty Model

- ξ_i: latent 'frailty' for ith pair
- $(T_{i1}, \ldots, T_{iK})$ independent given ξ_i
- Hazard for T_{ik} given ξ_i

$$\xi_i \alpha_{0k}(t)$$

- ξ belong to parametric family $g_\gamma(\cdot)$
  γ*: dependence parameter*
- Induces joint model for failures

Joint Survival Function

- $p_\gamma(\cdot)$ and $q_\gamma(\cdot)$: Laplace & it's inverse of $g(\cdot)$
- Let $\mathcal{S}_k(\cdot)$ be marginal survivor function

$$\mathrm{pr}(T_1 > t_1, \ldots, T_K > t_K) = p_\gamma \left[\sum_{k=1}^{K} q_\gamma\{\mathcal{S}_k(t_k)\} \right]$$

- Joint survival: function of $\mathcal{S}_k(\cdot)$ and γ
- Form of the function depends on $g(\cdot)$

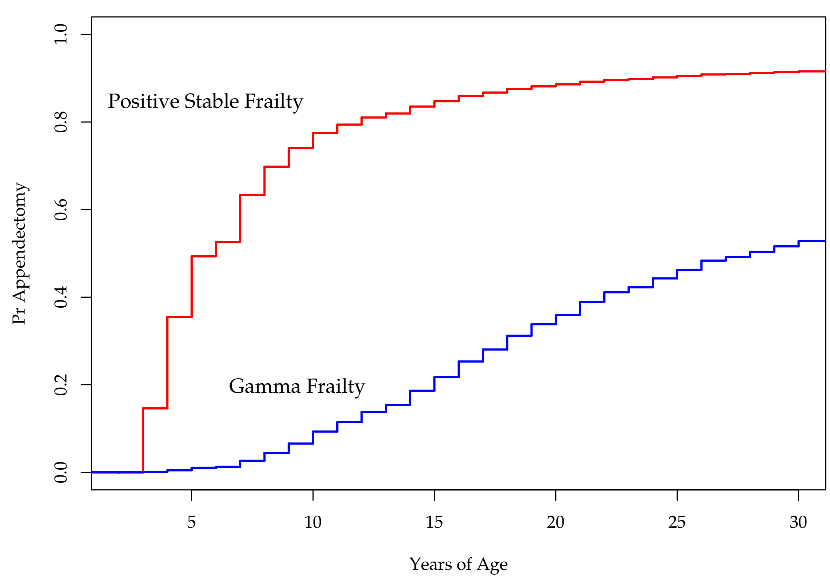
Why use a frailty model?

- Easy to specify joint model
  marginal specification and frailty distribution
- Parsimonious model
  flexible margins, parametric dependence
- Interpretable summaries of dependence
- Facilitates estimation of joint quantities
  e.g., conditional risk of failure

Australian Twin Study

- 1218 Female monozygotic twins
  age at appendectomy recorded
- One twin has appendectomy at age 3
- What is risk of co-twin?
  $\text{pr}(T_2 \leq t_2 | T_1 = 3, T_2 > 3), t_2 > 3$
- Data too sparse to answer non-parametrically
- Do different frailties give similar answers?

Conditional Risk of Appendectomy



Bivariate Association

- Frailty unobserved
- Choice of $g(\cdot)$ reflected by cross-ratio (CR) function.

$$\theta(t_1, t_2) := \frac{\lambda_1(t_1|T_2 = t_2)}{\lambda_1(t_1|T_2 \geq t_2)} := \frac{\lambda_2(t_2|T_1 = t_1)}{\lambda_2(t_2|T_1 \geq t_1)}$$

- Relative hazard at t_1
  given failure history of T_2 at t_2
- $\theta(\cdot, \cdot) \Longleftrightarrow g(\cdot)$
  Oakes (1989)

A Local Association Measure

	$\text{pr}(T_1 = t_1)$	$\text{pr}(T_1 > t_1)$	
$\text{pr}(T_2 = t_2)$	a	b	$a + b$
$\text{pr}(T_2 > t_2)$	c	d	
	$a + c$		n

$$S(t_1, t_2) = \text{pr}(T_1 > t_1, T_2 > t_2)$$

$$\theta(t_1, t_2) = \frac{S(dt_1, dt_2) S(t_1^-, t_2^-)}{S(t_1^-, dt_2) S(dt_1, t_2^-)}$$

$$\theta(t_1, t_2) = \frac{a/(a + b)}{(a + c)/n}$$

Selecting a Frailty Model

- Gamma: Constant CR function
- Positive Stable: CR markedly decreasing in time
  converges to 1
- Inverse Gaussian: CR decreasing in time
  converges to some limit
- Choice of model can lead to different inference

Model Checking

- Most methods focused on gamma frailty
  Shih & Louis; Shih; Glidden
  key computations easy for gamma frailty
- These methods can test for lack-of-fit
- Graphics lack strong interpretation
- Methods don't lead to selection of better fit
- Fan et. (2000) estimated CR nonparametrically

Checking Methods

A useful method would....

- Be able to check any model
- Unify graphical/inferential methods
- Closely tied to cross-ratio functions

I'll describe such an approach.

Consider pairs then extend

Three Hazard Functions

- $\Lambda_{11}(\cdot,\cdot), \Lambda_{10}(\cdot,\cdot), \Lambda_{01}(\cdot,\cdot)$: bivariate hazards
- Hazards of double and single failures
- Together they specify joint survival function
- Estimates have Aalen-like structure
- Theory for hazards estimates developed
  Gill, van der Laan, Wellner (1995)
  consistency, weak convergence, bootstrap

Bivariate Hazard Functions

By definition of these functions

$$\theta(t_1, t_2) = \frac{\Lambda_{11}(dt_1, dt_2)}{\Lambda_{10}(dt_1, t_2^-)\Lambda_{01}(t_1^-, dt_2)}$$

thus

$$\Lambda_{11}(dt_1, dt_2) - \theta(t_1, t_2)\Lambda_{10}(dt_1, t_2^-)\Lambda_{01}(t_1^-, dt_2) = 0 \quad (1)$$

when $\theta(t_1, t_2)$ correctly specified

The Approach

- Fit desired frailty model to the data
  lots of ways to do this
- Based on model, estimate $\hat{\theta}(\cdot, \cdot)$
  shape determined by model
- Estimate $(\hat{\Lambda}_{11}, \hat{\Lambda}_{10}, \hat{\Lambda}_{01})$
  fully non-parametric
  Neslon-Aalen type form
- Then calculate residuals based on formula (1)

A Checking Process

- Residuals at observed failure times
- Residual at (t_1, t_2):

$$\hat{\Lambda}_{11}(dt_1, dt_2) - \hat{\theta}(t_1, t_2)\hat{\Lambda}_{10}(dt_1, t_2^-)\hat{\Lambda}_{01}(t_1^-, dt_2)$$

- Magnitude of residual related to CR function
  positive: CR locally underestimated
  negative: CR locally overestimated

The Residuals: Heuristically

$\{t_{1(1)}, \ldots, t_{1(D_1)}\}$: D_1 failures of T_1
$\{t_{2(1)}, \ldots, t_{2(D_2)}\}$: D_2 failures of T_2

At failure times, form the 2-by-2 table

	$T_1 = t_{1(i)}$	$T_1 > t_{1(i)}$	
$T_2 = t_{2(j)}$	Ψ_{ij}		d_{2ij}
$T_2 > t_{2(j)}$			
	d_{1ij}		R_{ij}

Ψ_{ij} : # of pairs: $T_1 = t_{1(i)}$ and $T_2 = t_{2(j)}$

Residual: $\Psi_{ij} - \hat{\theta}\left[t_{1(i)}, t_{2(j)}\right] \frac{d_{1ij}d_{2ij}}{R_{ij}}$

Similar Methods

- Hsu & Prentice: test of independence
  same residuals with $\theta(t_1, t_2) = 1$
- Fan, Prentice, Hsu (2000)
  used similar expression for
  non-parametric estimate of CR function
- Clayton (1978) used Ψ for estimation in gamma frailty model

The Checking Process

- The model-checking process:

$$\hat{F}(t_1, t_2) = \sqrt{n} \int_0^{t_2} \int_0^{t_1} W(s_1, s_2) \{ \hat{\Lambda}_{11}(ds_1, ds_2) -$$

$$\hat{\theta}(s_1, s_2) \hat{\Lambda}_{10}(ds_1, s_2^-) \hat{\Lambda}_{01}(s_1^-, ds_2) \}$$

- Process: cumulative sum of residuals
- Residuals proportional to the 2 by 2 table residuals
- $W(\cdot, \cdot)$: arbitrary weight function
  e.g., size of risk set, estimated surv fn.

Appealing Features

- Interpretable process
- $\hat{F}(t_1, t_2)$ increases, CR locally underestimated
- $\hat{F}(t_1, t_2)$ decreases, CR locally overestimated
- Any $\hat{\theta}(\cdot, \cdot)$ can be used
- Graphical and testing component
- Can use some existing theory!!
  Gill et. al (1995)

Covariates

- Marginal hazard for kth failure has structure
- $A_{ik}(dt) = \alpha_{0k}(t) \exp(\beta Z_{ik})$
- Z is a time-fixed covariate
- Under frailty model if margins known, then
- $[\exp\{-A_{i1}(T_{i1})\}, \exp\{-A_{i2}(T_{i2})\}]$
  uniform margins
  common CR function across clusters
- $\theta_\alpha(u_1, u_2)$ the common CR on unit square

Approach

- Estimate model: $(\hat{\gamma}, \hat{A}_0, \hat{\beta})$
- Specifies $\hat{\theta}_\alpha(\cdot, \cdot)$
- Create $\hat{U}_{ik} = \hat{A}_{ik}(X_{ik})$
- Use $(\hat{U}_{ik}, \delta_{ik})$ to estimate the bivariate hazards
- Check based on

$$\hat{F}^\alpha = \sqrt{n} \int \int W(u_1, u_2)\{\hat{\Lambda}_{11}^\alpha(du_1, du_2) -$$

$$\hat{\theta}_\alpha(u_1, u_2)\hat{\Lambda}_{10}^\alpha(du_1, u_2^-)\hat{\Lambda}_{01}^\alpha(u_1^-, du_2)\}$$

Arbitrary Cluster Sizes

- Failure vectors $(T_1, \ldots, T_K)$
- Assume: $\mathrm{pr}(T_j > t_1, T_k > t_2) = S(t_1, t_2)$
  identical margins, exchangeable structure
- Margins not identical:
  use prev. method to transform data
- Estimates of can be constructed of $(\hat{\Lambda}_{11}, \hat{\Lambda}_{10}, \hat{\Lambda}_{01})$
  in a Nelson-Aalen spirit
- Define a model-checking process using formula for F

Under Correct Frailty

- And regularity of $\sqrt{n}\{\hat{\theta}(t_1, t_2) - \theta_0(t_1, t_2)\}$
- $\hat{F}(t_1, t_2)$: approximately mean 0
- $\hat{F}(t_1, t_2) \Longrightarrow \mathcal{F}(t_1, t_2)$

  $\mathcal{F}(t_1, t_2)$ *mean 0, Gaussian process*
- $\{F^{\sharp}(t_1, t_2) - \hat{F}(t_1, t_2)\} \Longrightarrow \mathcal{F}(t_1, t_2)$

  $F^{\sharp}$*: bootstrapped process*
  bootstrap can approximate $\mathcal{F}(t_1, t_2)$
- Results can be shown for covariates and $K > 2$

How Proofs Work

- Repeated use of a delta method
- $\hat{F}$ is a smooth function of empiricals
- Empirical dealt with by empirical process theory
  done by Gill et al. (1995)
- Smooth: compactly differentiable
  - chain rule applies
  - transfers asymptotic theory
  - preserves the boostrap
- Gill (1989) lays out the master plan

Test for Fit

- Consider two simple summaries of $\hat{F}(\cdot,\cdot)$
- $Z_\tau = \hat{F}(\tau_1,\tau_2)/\hat{\text{SE}}\left\{\hat{F}(\tau_1,\tau_2)\right\}$
  compared to N(0,1)
  τ_1, τ_2 *some late time*
  bootstrap used to estimate SE
- $Q = \sup_t |\hat{F}(t_1,t_2)|$
  compare to $Q^\sharp = \sup |\{F^\sharp - \hat{F}\}|$

Gamma Frailty Simulations

- Generate gamma & positive stable frailty data
- Fit gamma frailty model
- 30% censoring
- $n = 50$, 100, 200 pairs
- Kendall's τ : 0.25,0.50,0.75
- 2,000 datasets, 200 bootstrap samples
- Empirical size of Z_τ and Q statistics

Size of Tests: Gamma Frailty

τ		$n = 50$	$n = 100$	$n = 200$
0.25	Z_τ	0.06	0.07	0.07
	Q	0.05	0.06	0.07
0.50	Z_τ	0.06	0.06	0.05
	Q	0.04	0.04	0.05
0.75	Z_τ	0.03	0.05	0.04
	Q	0.02	0.03	0.04

Power to Reject Gamma Frailty

τ		$n = 50$	$n = 100$	$n = 200$
0.25	Z_τ	0.34	0.57	0.84
	Q	0.35	0.60	0.89
0.50	Z_τ	0.74	0.97	1.00
	Q	0.73	0.96	1.00
0.75	Z_τ	0.81	1.00	1.00
	Q	0.70	1.00	1.00

Positive Stable Frailty Simulations

- Generate positive stable & gamma frailty data
- Fit positive stable frailty model
- 30% censoring
- $n = 50$, 100, 200 pairs
- Kendall's τ : 0.25,0.50,0.75
- 2,000 datasets, 200 bootstrap samples
- Empirical size of Z_τ and Q statistics

Size of Tests: Positive Stable Frailty

τ		$n = 50$	$n = 100$	$n = 200$
0.25	Z_τ	0.08	0.07	0.06
	Q	0.06	0.04	0.06
0.50	Z_τ	0.07	0.07	0.06
	Q	0.04	0.05	0.06
0.75	Z_τ	0.09	0.06	0.06
	Q	0.03	0.06	0.05

Power: Positive Stable Frailty

τ		$n = 50$	$n = 100$	$n = 200$
0.25	Z_τ	0.38	0.66	0.92
	Q	0.20	0.49	0.83
0.50	Z_τ	0.88	1.00	1.00
	Q	0.66	0.97	1.00
0.75	Z_τ	0.97	1.00	1.00
	Q	0.85	1.00	1.00

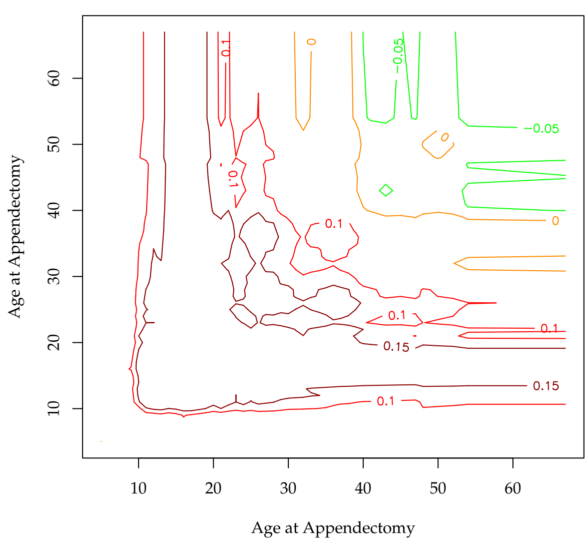
Australian Twin Study

- 1218 Australian female monozygotic twins
- Age at appendectomy recorded
  censored at current age
- Duffy et al (1990) *Am J Hum Genet*
- Strong clustering
  higher in MZ than DZ twins
- Gamma Frailty fit: $\hat{\theta}(t_1, t_2) = 2.83$
  significant clustering

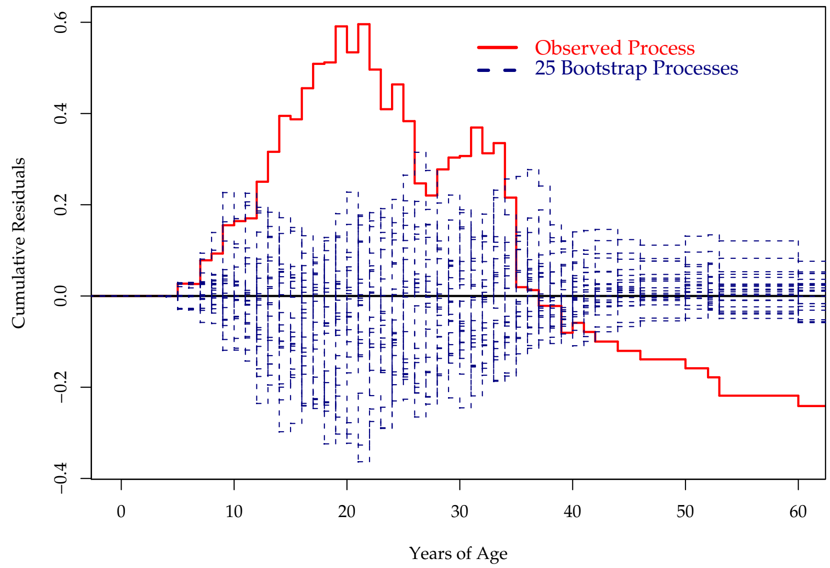
Constant CR function?

- Non-constant CR function suggested
  Fan et al, 2000; Kooperberg
- Glidden (1999): martingale residual method
  non-significant, $p = 0.12$
- Calculated test: 10,000 bootstrap samples
  Z_τ: *p=0.0004, Q: p=0.0026*
- Positive stable fits reasonably:
  Q: p=0.38

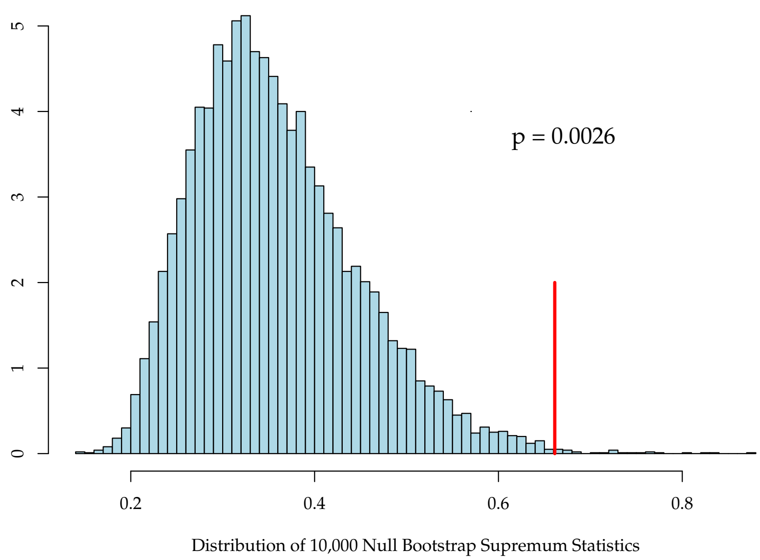
The Process $\hat{F}(\cdot,\cdot)$



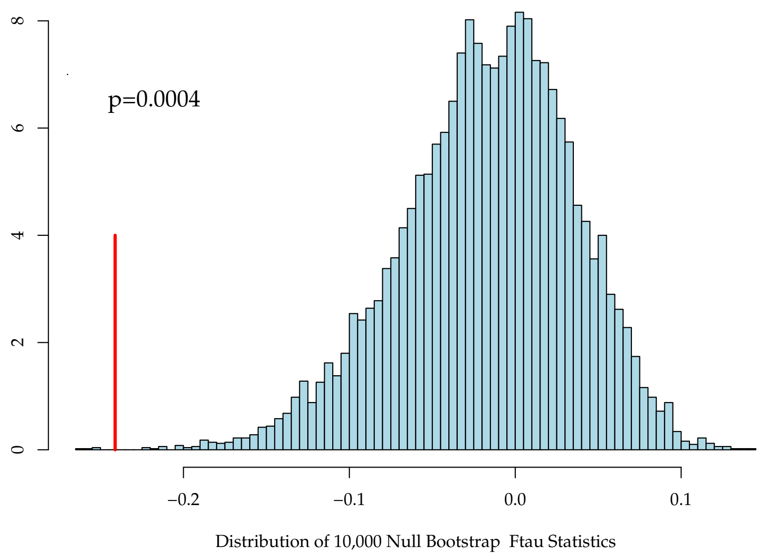
Observed and Null Processes



Observed and Null Q Statistic



Observed and Null Z_τ Statistic



Results

- Interpretable, general method
- Properties of F dependent on $\hat{\theta}(\cdot, \cdot)$
  no completely general theory
- Bootstrap works well
  well-attuned to approximate null distn
- 3-D plots perhaps overwhelming
  2-D perhaps more manageable