General copula-checking strategies for bivariate failures

Dave Glidden, UCSF

June 28, 2004

WNAR

Background

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

Bivariate Failure Time

- Data sampled in clusters (families)
  $i = 1, \ldots, n$
- Clusters contain paired failures (T_1, T_2)
  can be extended beyond pairs
- Example: Ages of appendectomy
  Duffy, Australian Twin Study (1990)
- How to describe, model dependence?

Bivariate Failure Time

- ξ_i: latent 'frailty' for ith pair
- (T_{i1}, T_{i2}) independent given ξ_i
- Hazard for T_{ij} given ξ_i

$$\xi_i \lambda_{0j}(t)$$

- ξ belong to parametric family $g(\cdot;\beta)$
- Choice of $g(\cdot)$ affects 'dependence'
  gamma, positive stable, inverse Gaussian

Association in Frailty Models

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

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

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

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

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

A Local Association Measure

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

- $\theta(t_1, t_2) := an/(a+b)(a+c)$
- $H_{11}(dt_1, dt_2) = a/n$
- $H_{10}(dt_1, t_2) = (a+c)/n$
- $H_{01}(t_1, dt_2) = (a+b)/n$
- $\theta(t_1, t_2) := H_{11}(dt_1, dt_2)/H_{10}(dt_1, t_2)H_{01}(t_1, dt_2)$

Overview

- Has Mantel-Haenszel feel
- Cumulative sum of residuals at (t_1, t_2)
- Residuals based on 2 by 2 table
- Residuals $\propto \hat{\theta}_{np}(t_1, t_2)$ - $\hat{\theta}_p(t_1, t_2)$

  non-model-based - model based

  observed - expected

The Residuals

$\{t_{1(1)}, \ldots, t_{1(d_1)}\}$: 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

	$T_1 = t_{1(i)}$	$T_1 > t_{1(i)}$	
$T_2 = t_{2(j)}$	Δ_{ij}		$n_{2(j)}$
$T_2 > t_{2(j)}$			
	$n_{1(i)}$		n_{ij}

A Checking Process

- $\hat{\theta}_p(\cdot,\cdot)$: estimated CR function
- The model-checking process:

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

$$\hat{\theta}_p(s_1,s_2)\hat{H}_{10}(ds_1,s_2)\hat{H}_{01}(s_1,ds_2)\}$$

- Process: cumulative sum of residuals
- Residual at (s_1,s_2): $d\hat{H}_{11} - \hat{\theta}_p(s_1,s_2)d\hat{H}_{10}d\hat{H}_{01}$
  Proportional to the 2 by 2 table residuals
- $W_n(\cdot,\cdot)$: arbitrary weight function
  e.g., size of risk set, estimated surv fn.

Three hazard functions

- $H_{11}(\cdot,\cdot), H_{10}(\cdot,\cdot), H_{01}(\cdot,\cdot)$: 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

Appealing Features

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

Under Correct Frailty

- & Regularity on $\sqrt{n}\{\hat{\theta}_p(t_1, t_2) - \theta_0(t_1, t_2)\}$
- $F(t_1, t_2)$: approximately mean 0
- $F(t_1, t_2) \Longrightarrow \mathcal{F}(t_1, t_2)$

  $\mathcal{F}(t_1, t_2)$ *mean 0, Gaussian process*

- $\{F^\sharp(t_1, t_2) - F(t_1, t_2)\} \Longrightarrow \mathcal{F}(t_1, t_2)$

  $F^\sharp$*: bootstrapped process*

  bootstrap can approximate $\mathcal{F}(t_1, t_2)$

Test for Fit

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

Simulations: Null

- Generate gamma frailty data
- Fit gamma frailty model
- 50% type I censoring
- 50, 100, 200 pairs
- Kendall's τ : 0.25,0.50,0.75
- 2,000 datasets, 200 bootstrap samples
- Empirical size of F_τ and Q for gamma frailty

Size of Tests: Gamma Frailty

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

Simulations: Alternative

- Generate positive stable frailty data
- Fit gamma frailty model
- 50% type I censoring
- $n = 50$, 100, 200 pairs
- Kendall's τ : 0.25,0.50,0.75
- 2,000 datasets, 200 bootstrap samples
- Estimated power based on F_τ and Q statistics

Power to Reject Gamma Frailty

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

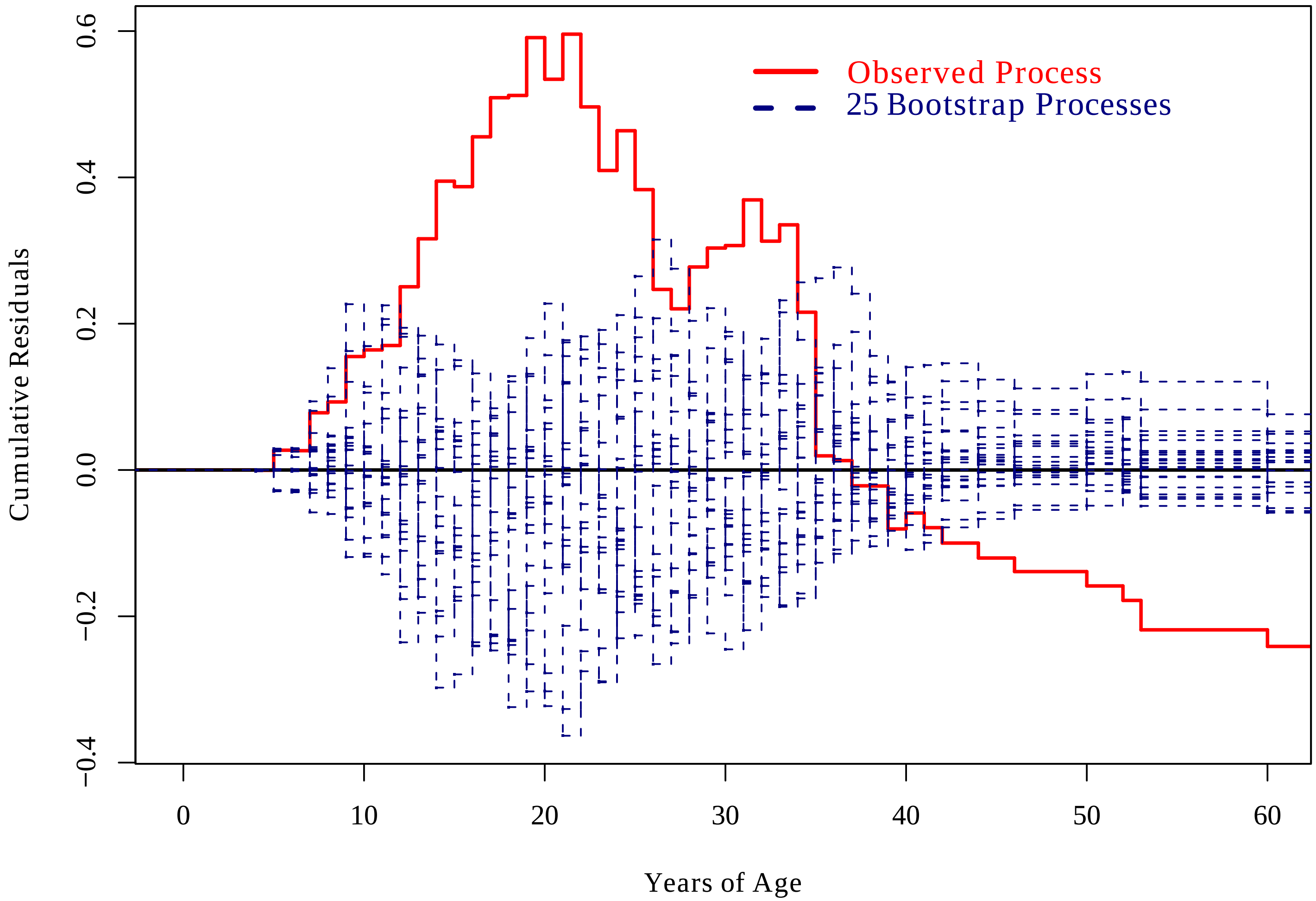
Australian Twin Study

- 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: $\hat{\theta}_p(t_1, t_2) = 2.83$
  significant clustering

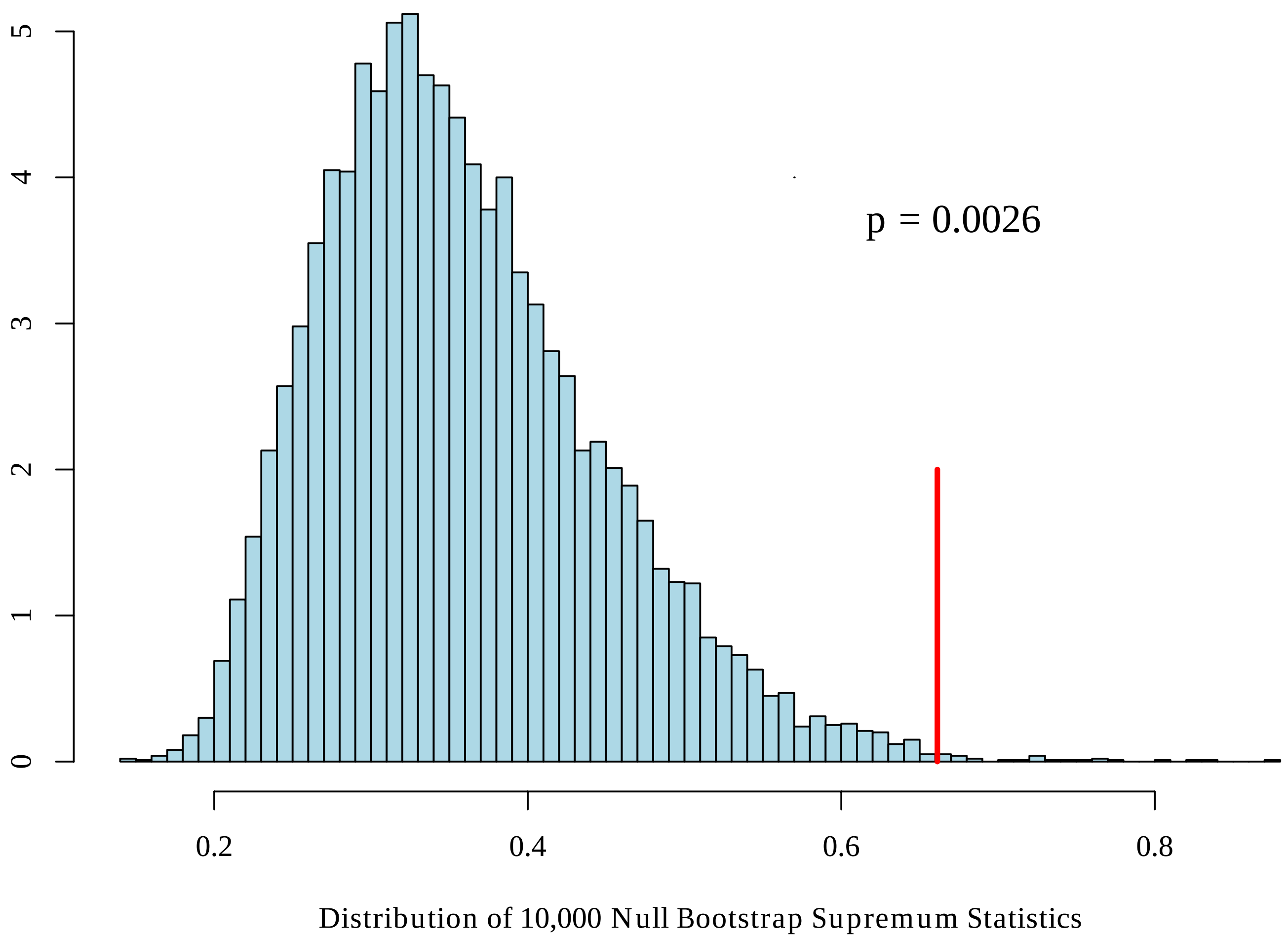
Constant CR function?

- Non-constant CR function suggested
  Fan et al, 2000; Kooperberg
- Glidden (1999): martingale residual method
  non-significant, $p = 0.12$
- Applied method
  unequivocal rejection of gamma frailty
- 3-D plot quite overwhelming
  can graph $F(t,t), t \in [0, \min(\tau_1, \tau_2)]$

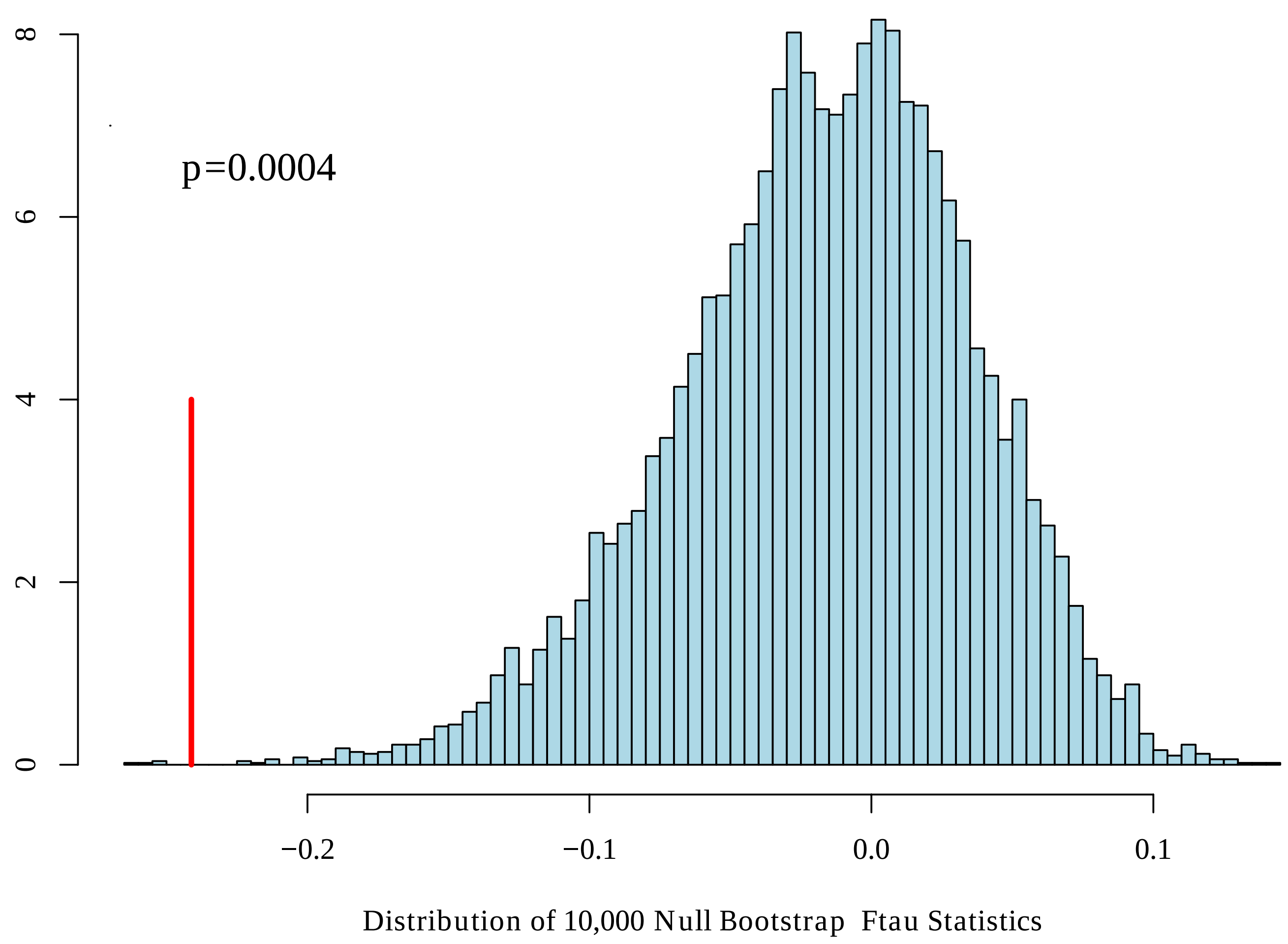
Observed and Null Processes



Observed and Null Q Statistic



Observed and Null F_τ Statistic



Results

- Properties of F dependent on $\hat{\theta}_p$
  no completely general theory
- Bootstrap work well
  well-attuned to approximate null distn
- Good properties, good power
- 3-D plots unwieldy
  2-D more informative, manageable
- F_τ powerful for monotone alternatives
  e.g. gamma vs. positive stable