Stat Seminar 10/26/06

‘General position results on uniqueness of nonrandomized Group-sequential decision procedures in Clinical Trials

Eric Slud, UMCP

OUTLINE

I. Two-Sample Clinical Trial Statistics

- A. Large-sample background
- B. Loss-functions & constraints on actions

II. Decision Theory Formulation

III. General Two-look Optimized Plans

- A. Reductions & results
- B. Computed Boundary – Example
- C. Role of uniqueness & nonrandomized solutions

IV. Random Perturbations of Loss Functions

- A. General result on a.s. unique minima
- B. Proof Sketch: a.s. unique optimal plans

Joint work with Eric Leifer, NHLBI.

Two-Sample Clinical Trial Statistics

Data format : $(E_i\,,\,T_i^*\,,\,\Delta_i^*\,,\,Z_i\,,\quad i=1,\ldots,N_A(\tau)\,)$
for analysis at times t_*.

E_i entry-times, N_A arrival-counting, τ accrual-horizon

X_i failure time, C_i indep. right-cens., Z_i trt. gp.

(X_i, C_i cond. indep. given Z_i & strat. variable V_i)

$T_i^* = X_i \wedge C_i \wedge (t_* - E_i)\,,\quad \Delta_i = I_{[X_i \le C_i \wedge (t_* - E_i)]})$

PROBLEM: test $H_0:\ S_{X|Z}(t|z) \equiv S_X(t)\,,\ z = 0,1$
with multiple interim looks & experimentwise validity.

TEST STATISTIC : for look at t_*, define

$Y_z^*(s) = \sum_j I_{[Z_j = z,\, T_j^* \ge s]}\,,\quad Y^*(s) = Y_1^*(s) + Y_0^*(s)$ *at-risk*

$$M(t_*) = -\sum_i \int K(s, \hat{S}_X(s)) \{ Z_i - \frac{Y_1^*(T_i^* \wedge s)}{Y^*(T_i^* \wedge s)} \}\, \Delta_i^*\, dI_{[T_i \le s]}$$

asympt. indep. incr., with estimated variance $\hat{V}(t_*)$ =

$$\sum_i \int K^2(s, \hat{S}_X(s)) \{ \frac{Y_1^*(T_i^* \wedge s) Y_0^*(T_i^* \wedge s)}{Y^*(T_i^* \wedge s)} \}\, \Delta_i^*\, dI_{[T_i \le s]}$$

Reject if $\frac{M(t_*)}{\sqrt{\hat{V}(t_*)}} \ge b(t_*)$, **Accept** if $\le a(t_*)$

Abstracted Asymptotic Problem

Most methods rely on asymptotically Gaussian time-indexed statistic-numerator $n^{-1/2} M(t)$ with indep. incr.'s, and variance function $V(t)$ to be estimated in real or information time.

$V(t)$ in H_0 is functional of $(S_X, S_{C|Z}, \Lambda_A \equiv E(N_A))$

$$\int K^2(S_X(s)) \frac{p(1-p)\Lambda_A(t-s)S_{C|Z}(s|1)S_{C|Z}(s|0)}{\Lambda_A(t)(p(S_{C|Z}(s|1)+(1-p)S_{C|Z}(s|0))} dF_X(s)$$

Control parameters: At each $t = t_k$, can choose t_{k+1}, a_{k+1}, b_{k+1}

PROBLEM: To optimize expected Loss or Cost over times and cutoffs while maintaining overall nominal significance level.

Main Computational Method of optimizing boundaries is parametric search for parametric boundary classes, or **backward induction**.

Decision-Theory Ingredients

Actions: look-times t_{k*} and boundaries $b_{k*} = b(t_*)$ for $W/\hat{V}^{1/2}$ — for now consider only upper rejection boundary

Prior: $\pi(\vartheta)$ density for group-difference log hazard ratio parameter ϑ (within semiparametric model).

Losses: costs of experimentation $c_1(t, \vartheta)$, wrong decision $c_2(\vartheta)$, late correct decision $c_3(t, \vartheta)$,

these loss elements introduced in Leifer (2000) thesis. Costs are economic within-trial, ethical within-trial, and economic after-trial.

In 2-Look Trials, adaptive actions are:

- t_1 constant depending on prior & losses
- Followup time $t_2 = t_2(W(t_1), t_1)$ which if $= t_1$ indicates imediate Reject or Accept, and
- final boundary: reject iff $W(t_2) \geq b(t_2, t_1, W(t_1))$

More on Decision Theory Problem

(I) state-of-nature parameter $\vartheta \in \mathbf{R}$, *target* parameter $z = \gamma(\vartheta)$ $(= \vartheta$ in estimation, $I[\vartheta \le 0]$ in testing), and prior prob. π on Θ

(II) action-space

$$\mathcal{A} = \{(\underline{t}, a) \;:\; \underline{t} = (t_1, \ldots, t_K),$$
$$0 \le t_1 \le \cdots \le t_K \le T,\; a \in \gamma(\Theta)\}$$

(III) observable process $\mathbf{W} = (W(x),\;\; 0 \le x \le T)$ Wiener process with drift ϑ, and suff. stat. $W(\tau)$ for data up to stopping-time τ, and *auxiliary randomization* r.v. $U \sim \mathrm{Unif}[0,1]$ independent of $\mathbf{W}$

(IV) strategies consist of: increasing nonnegative stopping-time random variables $\underline{\tau} = (\tau_1, \ldots, \tau_K)$

$$0 \le \tau_1 \le \tau_2 \le \cdots \le \tau_K \le T \quad a.s.$$

satisfying 'causality' restrictions on measurability,

$$[\tau_j \le x] \in \sigma(\tau_i,\; i < j \;;\;\; W(u),\; u \le x \;\;;\; U),$$

and *terminal action* r.v. $a \in \gamma(\Theta)$ meas. wrt $(W(x),\; x \le t_K)$ and U

(V) a *loss-function* $L(\underline{t}, a, \vartheta, \lambda)$ depending on look-times $\underline{t}$ only as smooth fcn of terminal time t_K, with Lagrange multipliers $\lambda_0,\ \lambda_1$

$$L(\underline{t}, a, \vartheta, \lambda) \ = \ C(t_K, a, z) + \lambda_0\, a\, \delta_{\vartheta,0} + \lambda_1\, (1-a)\, \delta_{\vartheta,\vartheta_1}$$

where $z = I_{[\vartheta \le 0]}$, $\vartheta_1 > 0$ and

$$C(t, a, z) \ = \ c_1(t, \vartheta) \ + \ I_{[z \ne a]}\, c_2(\vartheta) \ + \ I_{[z=a]}\, c_3(t, \vartheta)$$

and $c_j(\cdot, \vartheta) \nearrow$, $c_3(t, \vartheta) \le c_2(\vartheta)$.

(VI) *risk function* to be minimized over $d = (\underline{\tau}, a)$ is

$$r(d, \lambda) \ = \ \int_\Theta E_\vartheta\, L(\underline{\tau}, \alpha, \vartheta, \lambda)\, d\pi(\vartheta) \qquad (1)$$

NOTE: For terminal time $\tau_K = s$ and data history $(W(u),\ u \le s)$, there exists unique optimal terminal action $I_{[W(s) \ge w(s)]}$, with $w(\cdot) \ = \ w(\cdot, \lambda_0, \lambda_1)$ uniquely determined by

$$\int e^{\theta w(y) - \theta^2 y/2}\, g_1(y, \vartheta, \lambda_0, \lambda_1)\, d\pi(\theta) \ = \ 0$$

where

$$g_1(y) = (c_2(\theta) - c_3(y, \theta))\, (2 I_{[\theta \le 0]} - 1) + \frac{\lambda_0}{\pi_0}\, I_{[\theta = 0]} - \frac{\lambda_1}{\pi_1}\, I_{[\theta = \theta_1]}$$

Optimizing the Decision Rule

We know that at the final look-time, the best decision is to `Reject` when $W(s) \geq w(s)$, w computed uniquely from s, λ_0, λ_1 .

Now let t_1 be fixed and imagine $W(t_1)/\sqrt{t_1} = x$ observed. Can express conditional expected final loss at time $t_1 + s$ in form

$$r_2(t_1, x, s) \equiv \int R(t_1, x, s, \vartheta)\, d\pi(\vartheta)$$

$$= \int e^{\theta x \sqrt{t_1} - \theta^2 t_1/2} \{a_0(t_1 + s, \vartheta) + a_1(t_1 + s, \vartheta) \cdot$$

$$\left(1 - \Phi\left(\frac{w(t_1 + s) - x\sqrt{t_1}}{\sqrt{s}} - \vartheta\sqrt{s}\right)\right) \}\, d\pi(\theta)$$

where

$$a_0(t, \vartheta) = c_1(t, \vartheta) + c_2(\vartheta)\, I_{[\vartheta > 0]} + \frac{\lambda_1}{\pi_1}\, I_{[\theta = \theta_1]} + c_3(t, \vartheta)\, I_{[\vartheta \leq 0]}$$

Backward Induction Idea: if we can uniquely optimize $r_2(t_1, x, s)$ in $s = s(t_1, x)$ uniquely ($a.e.\ x$), and then choose t_1 uniquely to minimize

$$\int \int R(t_1, x,\ s(t_1, x),\ \vartheta) \frac{1}{\sqrt{2\pi}}\, e^{-(x - \vartheta\sqrt{t_1})^2/2}\, dx\, d\pi(\vartheta)$$

Optimizing, Continued

then we would have specified the unique nonrandomized optimal decision rule through the constant t_1 and functions $s(t_1, x)$, $w(t_1 + s)$.

But for these general loss functions there is no hope of explicit formulas, and even showing unique optima is complicated by the form of optimal s which will generally be 0 on the complement of some interval.

This behavior is unavoidable, reflecting the necessity to stop after one look at the data at time t_1 *if the observed statistic value* $W(t_1)$ *is too extreme !*

We show this next is a computed example.

Optimal Boundary in 2-Look Example

Data: $W(t) = B(t_i) + \vartheta t_i, \quad i = 1,2 \quad t_1$ is fixed in advance, continuation-time $t_2 - t_1 \geq 0$ is chosen as function of $W(t_1)$.

Loss for stopping at τ with Rejection indicator z :

$$c_1(\tau,\vartheta) + c_3(\tau,\vartheta) + z\,(c_2(\vartheta) - c_3(\tau,\vartheta))\,(2I_{[\vartheta \leq 0]} - 1)$$

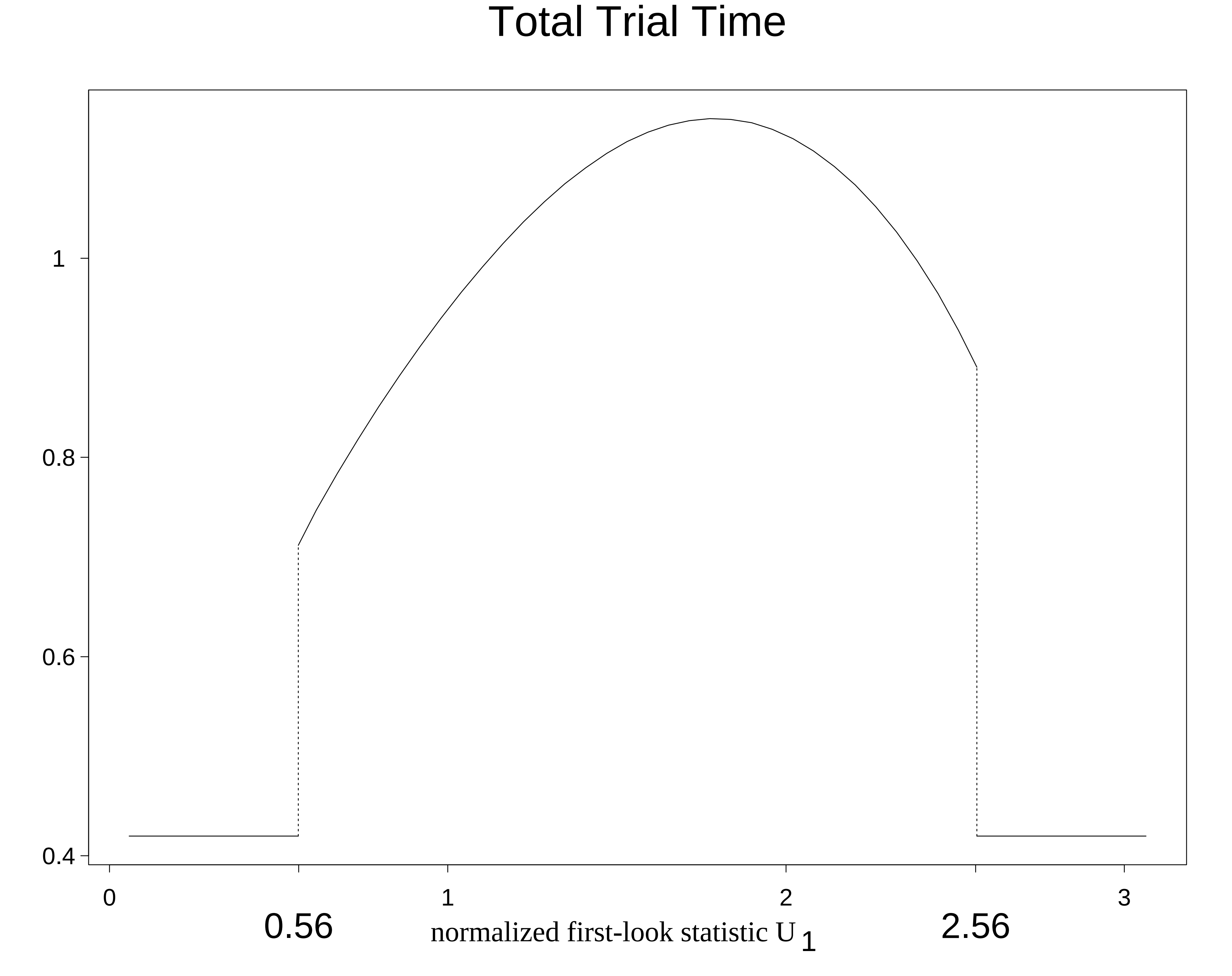
Problem to find min-risk test under prior $\pi(d\vartheta)$, with sig. level $\leq \alpha$ and type II error at $\vartheta_1 \leq \beta$.

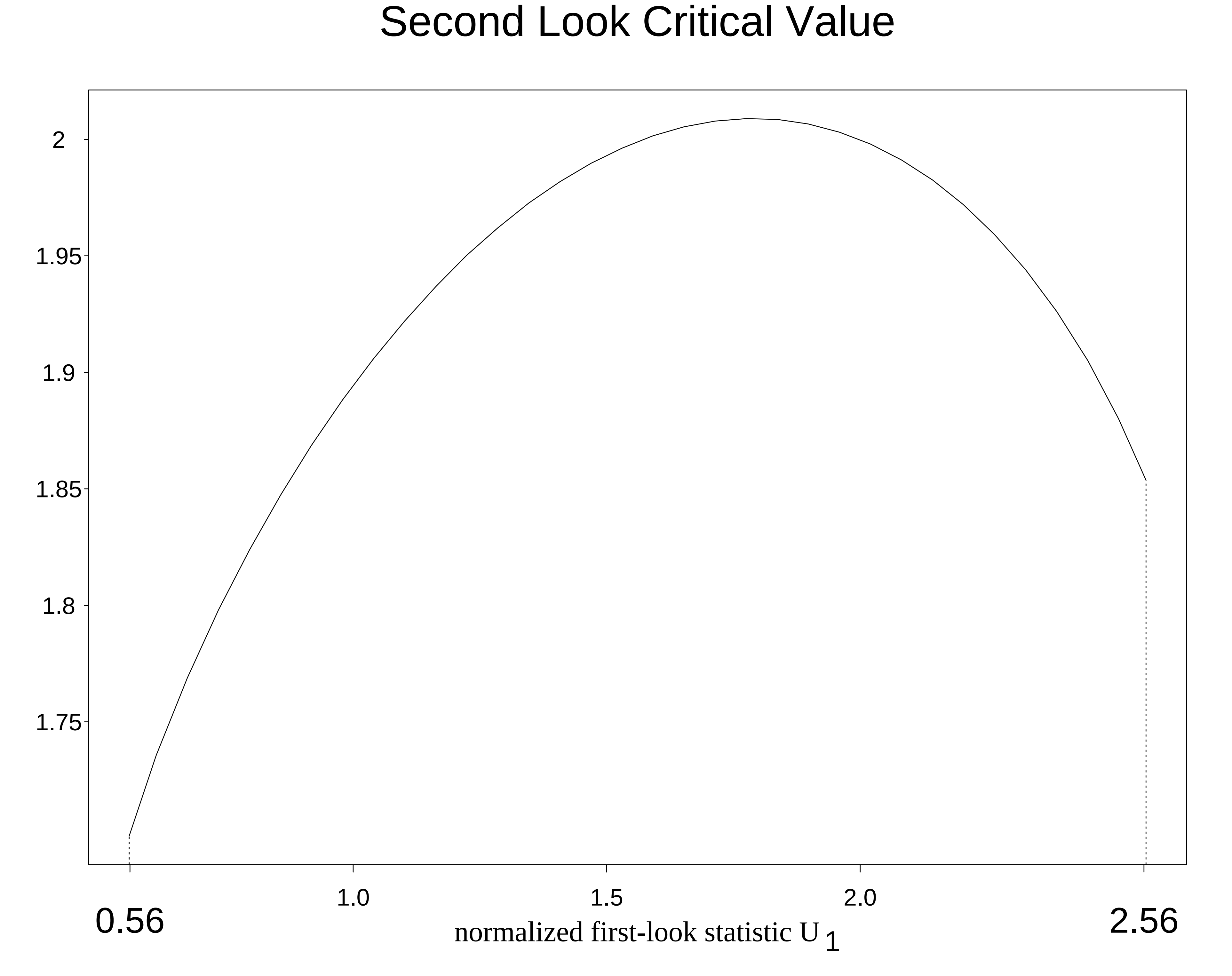
Under regularity conditions on loss elements (*piece-wise smoothness*, $c_2 \geq c_3$, $c_1 \nearrow \infty$) and prior $\pi(d\vartheta)$ assigning positive mass to neighborhoods of 0, $\vartheta_1 > 0$:

can show that optimal procedures are nonrandomized (w.p.1 after smallrandom perturbation of c_1) and unique, rejecting for $W(t_2) \geq b_2(t_1, t_2, W(t_1))\sqrt{V(t_2)}$.

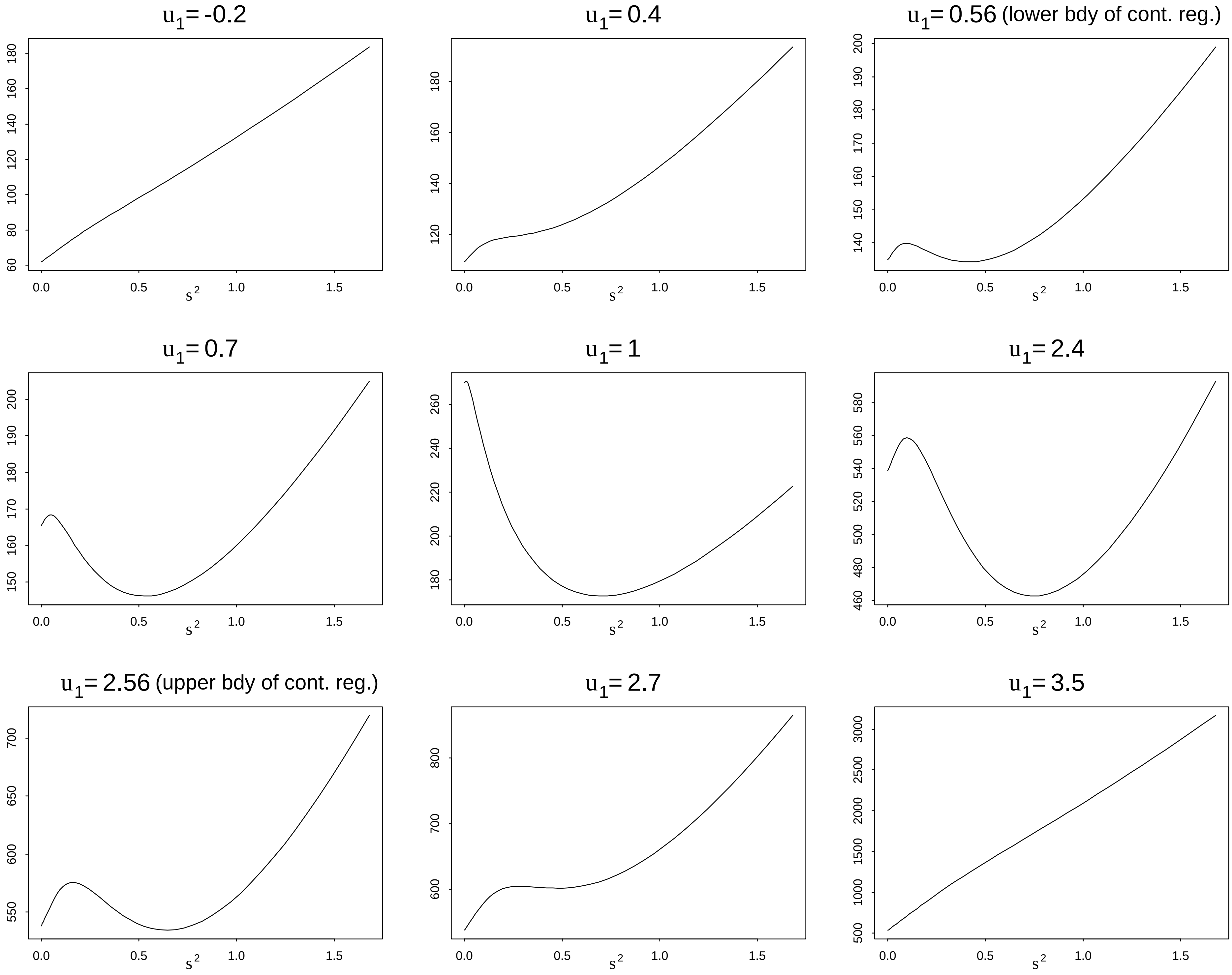
Example. $\alpha = .025$, $\beta = .1$, $\vartheta_1 = log(1.5)$, time scaled so $\tau_{fix} = 1$. Optimized $t_1 = .42 \cdot \tau_{fix}$.

e^ϑ = hazard ratio	0.9	1.0	1.25	1.5	1.75
$1.51 \cdot \pi(\{\vartheta\})$	0.2	1.0	0.2	0.1	0.01
$c_1(t,\vartheta)$	t	t	t	t	t
$c_2(\vartheta)$	200	100	50	250	500


Total Trial Time
1
0.8
0.6
0.4
0
1
2
3
0.56
normalized first-look statistic U 1
2.56



Second Look Critical Value
2
1.95
1.9
1.85
1.8
1.75
1.0
1.5
2.0
0.56
2.56
normalized first-look statistic U 1



u1= -0.2
u1= 0.4
u1= 0.56 (lower bdy of cont. reg.)
u1= 0.7
u1= 1
u1= 2.4
u1= 2.56 (upper bdy of cont. reg.)
u1= 2.7
u1= 3.5
s 2
0.0
0.5
1.0
1.5

Lemma on Unique Random-Function Minima

Lemma 1 *(extending Bulinskaya 1961) Let* K *be a fixed compact interval in* $\mathbf{R}$*, and suppose that* g_1, g_2 *are fixed smooth real-valued functions on* K*, with* $\inf_K g_2 > 0$*. Let* $\zeta(\cdot) = \zeta(\cdot, \omega)$ *be a random process on* K *with sample paths a.s. in* $\mathcal{C}^2(K)$ *such that for all* $\delta > 0$*, the joint density of* $(\zeta(s_1), \zeta'(s_1), \zeta(s_2), \zeta'(s_2))$ *(with respect to Lebesgue measure, on* $\mathbf{R}^4$*) is uniformly bounded, uniformly over all* $s_1, s_2 \in K$ *for which* $|s_1 - s_2| \geq \delta$*. Then*

(i) With probability 1 (ω*), there do not exist distinct elements* $s_1, s_2 \in K$ *such that the sample path of the function* $\rho(s) \equiv g_1(s) + g_2(s)\zeta(s)$ *satisfies* $\rho'(s_1) = \rho'(s_2) = 0$ *along with* $\rho(s_1) = \rho(s_2)$*, and*

(ii) Under the assumptions above, with probability 1, the number of zeroes of ρ' *on* K *is finite, and* $\rho''(s) \neq 0$ *at all values of* s *for which* $\rho'(s) = 0$*.*

Corollary. Analogous result for minima wrt s with functions $g_j(t_1, x, s)$ holds a.e. in (t_1, x, ω).

Via Implicit Function Theorem for minimizer $s = s(t_1, x)$ find s a.e. smooth in its arguments, a.s. for argument ω .

Further Steps to Establish a.s. unique rule

Want to substitute $s = s(t_1, x)$ and integrate out standardized variable $x = W(t_1)/\sqrt{t_1}$ to get smooth time-t_1 risk $r_1(t_1)$ **except** that this function s is only a.e. defined and smooth.

Introduce technical assumption for small $\epsilon > 0$ that

$$t_2 \geq t_1 + \epsilon \qquad \text{whenever} \qquad t_2 > t_1$$

(No real loss in generality: can likely prove it directly.)

Remaining proof-step to prove $r_1(t_1)$ smooth: for $V \sim \text{Unif}[0, \epsilon)$ indep. of other data,

$$r_1(t_1) \; = \; \min \left(E(r_2(t_1, W(t_1)/\sqrt{t_1}, 0+) \right),$$

$$E(\; r_2(t_1 + V, \; \frac{W(t_1 + V)}{\sqrt{t_1 + V}}, s(t_1 + V, \; \frac{W(t_1 + V)}{\sqrt{t_1 + V}}) \,))$$

Last expectations with respect to measure on $(\vartheta, W(\cdot))$ given by $\int_A P_\vartheta(dW) d\pi(\vartheta)$.

Overall results hold a.s. (ω) for a.e. (λ_0, λ_1). But that seems to be enough.

References

For clinical-trial large sample theory:

Tsiatis (1982), Andersen & Gill (1982), Sellke & Siegmund (1983), Slud (1984)

For group-sequential boundaries:

Pocock (1977), O'Brien & Fleming (1979), Slud & Wei (1982), Lan & DeMets (1983) , Pampallona & Tsiatis (1994) plus others

For adaptive modifications:

Burman & Sondesson (2006, current **Biometrics**)

For decision-theoretic formulations:

Anscombe (1963), Colton (1963), Berger (1985 book), Siegmund (1985 book), Jennison (1987), Freedman & Spiegelhalter (1989), Leifer (2000, thesis)

For level-crossings theorems:

Bulinskaya (1962) cited in Cramer and Leadbetter (1967) *Stationary and Related Stochastic Processes* book