Estimation and Testing
with Current Status Data

Jon A. Wellner

University of Washington

- joint work with Moulinath Banerjee,
  University of Michigan
- Talk at Université Paul Sabatier, Toulouse III,
  Laboratoire de Statistique et Probabilités,
- *Email: jaw@stat.washington.edu*
  http: //www.stat.washington.edu/jaw/jaw.research.html

Outline

- Introduction: current status data

Outline

- Introduction: current status data
- Estimation of F, unconstrained (old, Ayer et al. 1956)

Outline

- Introduction: current status data
- Estimation of F, unconstrained (old, Ayer et al. 1956)
- Estimation of F, constrained (new)

Outline

- Introduction: current status data
- Estimation of F, unconstrained (old, Ayer et al. 1956)
- Estimation of F, constrained (new)
- The likelihood ratio test of $H : F(t_0) = \theta_0$

Outline

- Introduction: current status data
- Estimation of F, unconstrained (old, Ayer et al. 1956)
- Estimation of F, constrained (new)
- The likelihood ratio test of $H : F(t_0) = \theta_0$
- How big is “too big”? The limit Gaussian problem

Outline

- Introduction: current status data
- Estimation of F, unconstrained (old, Ayer et al. 1956)
- Estimation of F, constrained (new)
- The likelihood ratio test of $H : F(t_0) = \theta_0$
- How big is “too big”? The limit Gaussian problem
- Limiting Distribution of the likelihood ratio statistic under H

Outline

- Introduction: current status data
- Estimation of F, unconstrained (old, Ayer et al. 1956)
- Estimation of F, constrained (new)
- The likelihood ratio test of $H : F(t_0) = \theta_0$
- How big is "too big"? The limit Gaussian problem
- Limiting Distribution of the likelihood ratio statistic under H
- Confidence intervals for $F(t_0)$

Outline

- Introduction: current status data
- Estimation of F, unconstrained (old, Ayer et al. 1956)
- Estimation of F, constrained (new)
- The likelihood ratio test of $H : F(t_0) = \theta_0$
- How big is “too big”? The limit Gaussian problem
- Limiting Distribution of the likelihood ratio statistic under H
- Confidence intervals for $F(t_0)$
- Further problems

1. Introduction: current status data

- $X \sim F, Y \sim G, \quad X, Y$ independent

1. Introduction: current status data

- $X \sim F$, $Y \sim G$, $\quad X, Y$ independent
- We observe $(Y, 1\{X \leq Y\}) \equiv (Y, \Delta)$ with density

$$p_F(y, \delta) = F(y)^{\delta}(1 - F(y))^{1-\delta} g(y)$$

1. Introduction: current status data

- $X \sim F$, $Y \sim G$, $\quad X, Y$ independent
- We observe $(Y, 1\{X \le Y\}) \equiv (Y, \Delta)$ with density

$$p_F(y,\delta) = F(y)^{\delta}(1 - F(y))^{1-\delta} g(y)$$

- Suppose that (Y_i, Δ_i) are i.i.d. as (Y, Δ).

1. Introduction: current status data

- $X \sim F$, $Y \sim G$, $\quad X, Y$ independent
- We observe $(Y, 1\{X \leq Y\}) \equiv (Y, \Delta)$ with density

$$p_F(y, \delta) = F(y)^\delta (1 - F(y))^{1-\delta} g(y)$$

- Suppose that (Y_i, Δ_i) are i.i.d. as (Y, Δ).
- Likelihood:

$$L_n(F) = \prod_{i=1}^{n} F(Y_i)^{\Delta_i} (1 - F(Y_i))^{1-\Delta_i}$$

- Likelihood ratio test of $H : F(t_0) = \theta_0$

- Likelihood ratio test of $H : F(t_0) = \theta_0$
- The likelihood ratio statistic:

$$\lambda_n = \frac{\sup_F L_n(F)}{\sup_{F:F(t_0)=\theta_0} L_n(F)} = \frac{L_n(\hat{F}_n)}{L_n(\hat{F}_n^0)}.$$

2. Estimation of F: Nonparametric MLE

- MLE (Unconstrained) $\hat{F}_n(t) = \text{argmax}_F L_n(F)$

2. Estimation of F: Nonparametric MLE

- MLE (Unconstrained) $\hat{F}_n(t) = \text{argmax}_F L_n(F)$
- Another description: define

$$\mathbb{G}_n(t) = \frac{1}{n}\sum_{i=1}^{n} 1_{[Y_i \le t]}, \qquad \mathbb{V}_n(t) = \frac{1}{n}\sum_{i=1}^{n} \Delta_i 1_{[Y_i \le t]} \, .$$

Note that

$$\mathbb{G}_n(t) \rightarrow_{a.s.} G(t), \qquad \mathbb{V}_n(t) \rightarrow_{a.s.} \int_0^t F(y) dG(y) \equiv V(t) \, .$$

2. Estimation of F: Nonparametric MLE

- MLE (Unconstrained) $\hat{F}_n(t) = \text{argmax}_F L_n(F)$
- Another description: define

$$\mathbb{G}_n(t) = \frac{1}{n}\sum_{i=1}^{n} 1_{[Y_i \le t]}, \qquad \mathbb{V}_n(t) = \frac{1}{n}\sum_{i=1}^{n} \Delta_i 1_{[Y_i \le t]} \,.$$

Note that

$$\mathbb{G}_n(t) \rightarrow_{a.s.} G(t), \qquad \mathbb{V}_n(t) \rightarrow_{a.s.} \int_0^t F(y)dG(y) \equiv V(t)\,.$$

- Thus

$$\frac{dV}{dG}(t) = F(t)$$

- Partial sum diagram: Let $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(n)}$. The *partial sum diagram* $\mathcal{P} = \{P_i\}$ is given by

$$P_i = (\mathbb{G}_n(Y_{(i)}), \mathbb{V}_n(Y_{(i)})), \quad i = 1, \ldots, n\,.$$

- Partial sum diagram: Let $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(n)}$.
  The *partial sum diagram* $\mathcal{P} = \{P_i\}$ is given by

$$P_i = (\mathbb{G}_n(Y_{(i)}), \mathbb{V}_n(Y_{(i)})), \quad i = 1, \ldots, n\,.$$

- The Nonparametric MLE $\hat{F}_n$ of F is:

$$\hat{F}_n(Y_{(i)}) = \text{left derivative of the Greatest Convex Minorant of } \mathcal{P} \text{ at } Y_{(i)}\,.$$

- Partial sum diagram: Let $Y_{(1)} \le Y_{(2)} \le \cdots \le Y_{(n)}$. The *partial sum diagram* $\mathcal{P} = \{P_i\}$ is given by

$$P_i = (\mathbb{G}_n(Y_{(i)}), \mathbb{V}_n(Y_{(i)})), \quad i = 1, \ldots, n\,.$$

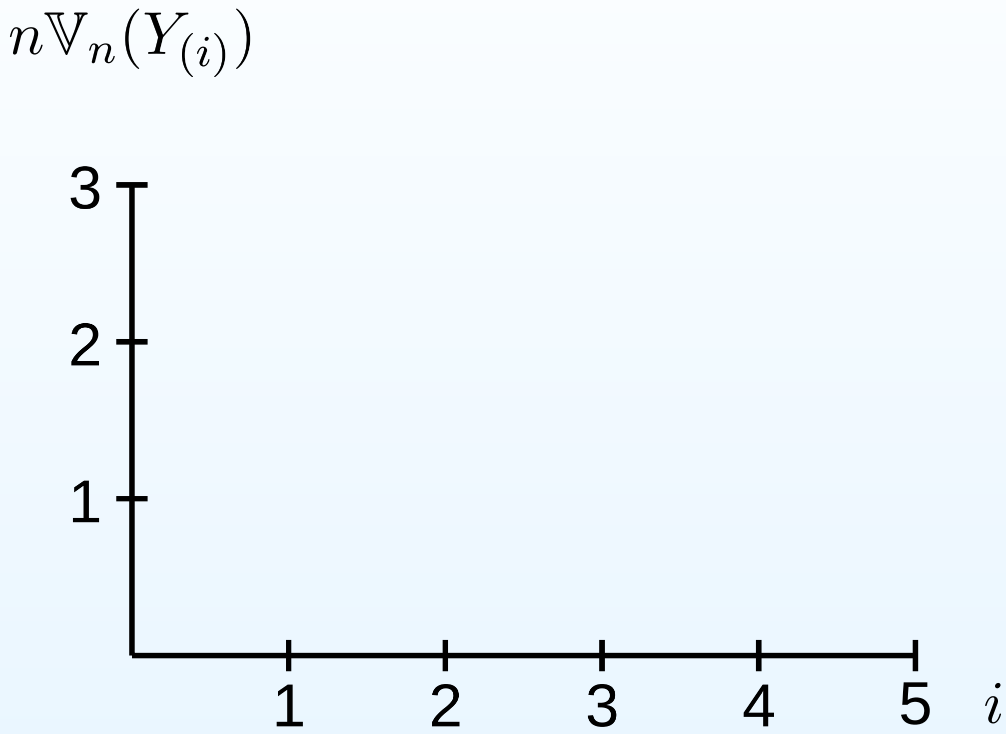
- The Nonparametric MLE $\hat{F}_n$ of F is:

$$\hat{F}_n(Y_{(i)}) = \text{left derivative of the Greatest Convex Minorant of } \mathcal{P} \text{ at } Y_{(i)}\,.$$

- Greatest Convex Minorant $=$ GCM

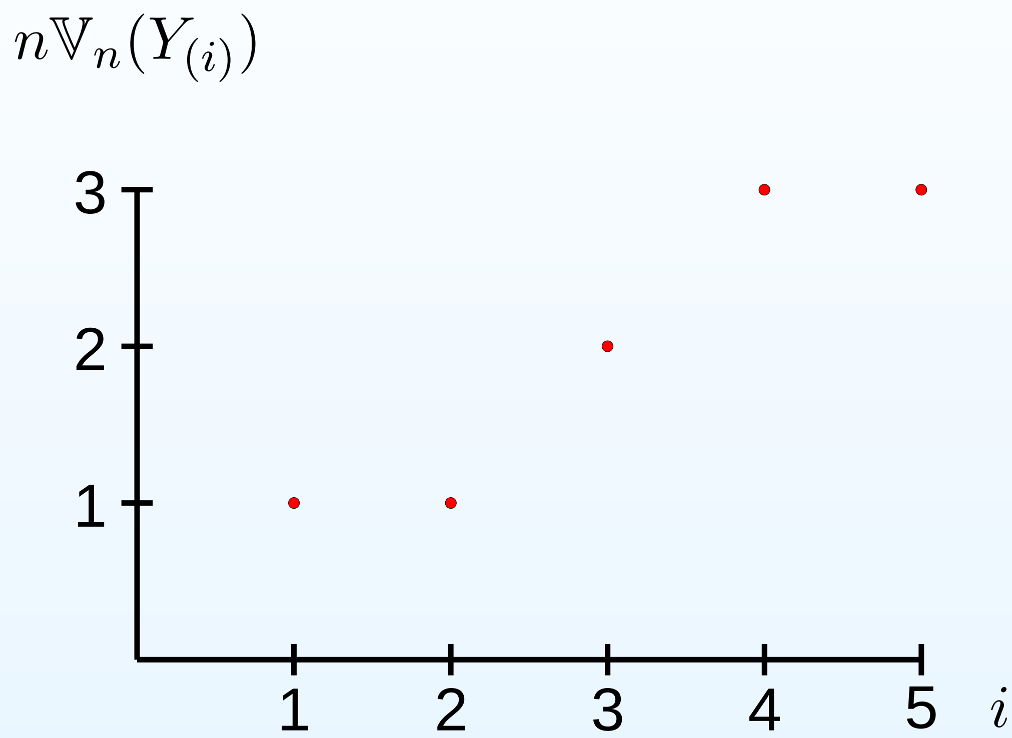
Cumulative Sum Diagram: Example $n = 5$

i	$\Delta_{(i)}$	$Y_{(i)}$
1	1	0.4
2	0	0.7
3	1	1.3
4	1	1.5
5	0	2.0



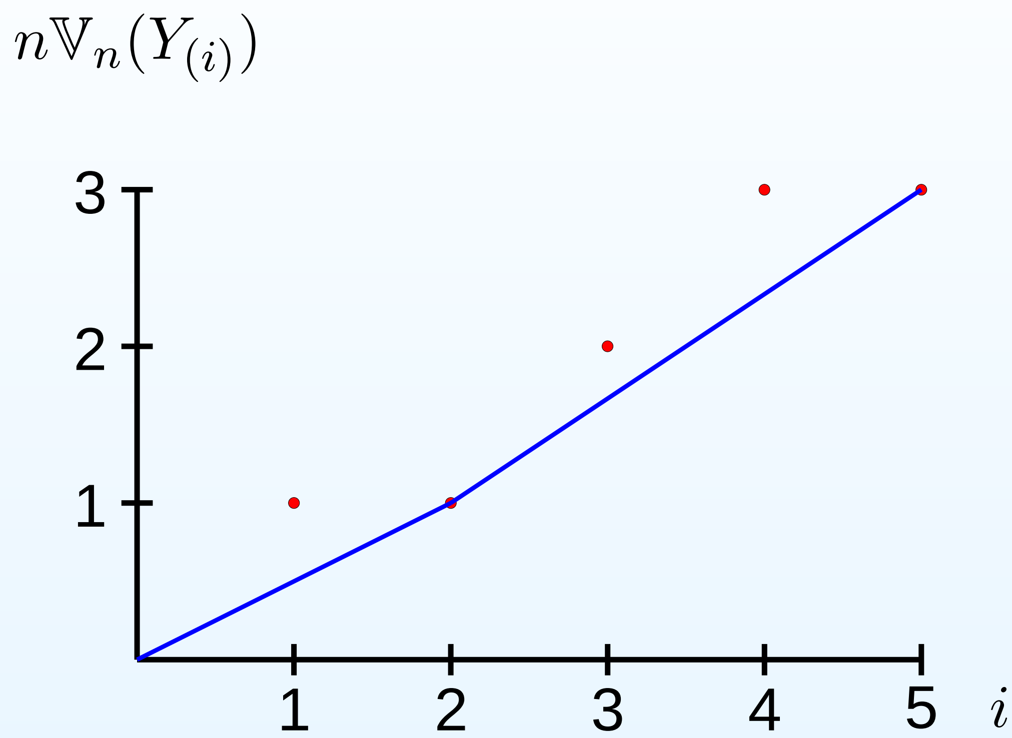
Cumulative Sum Diagram: Example $n = 5$

i	$\Delta_{(i)}$	$Y_{(i)}$
1	1	0.4
2	0	0.7
3	1	1.3
4	1	1.5
5	0	2.0



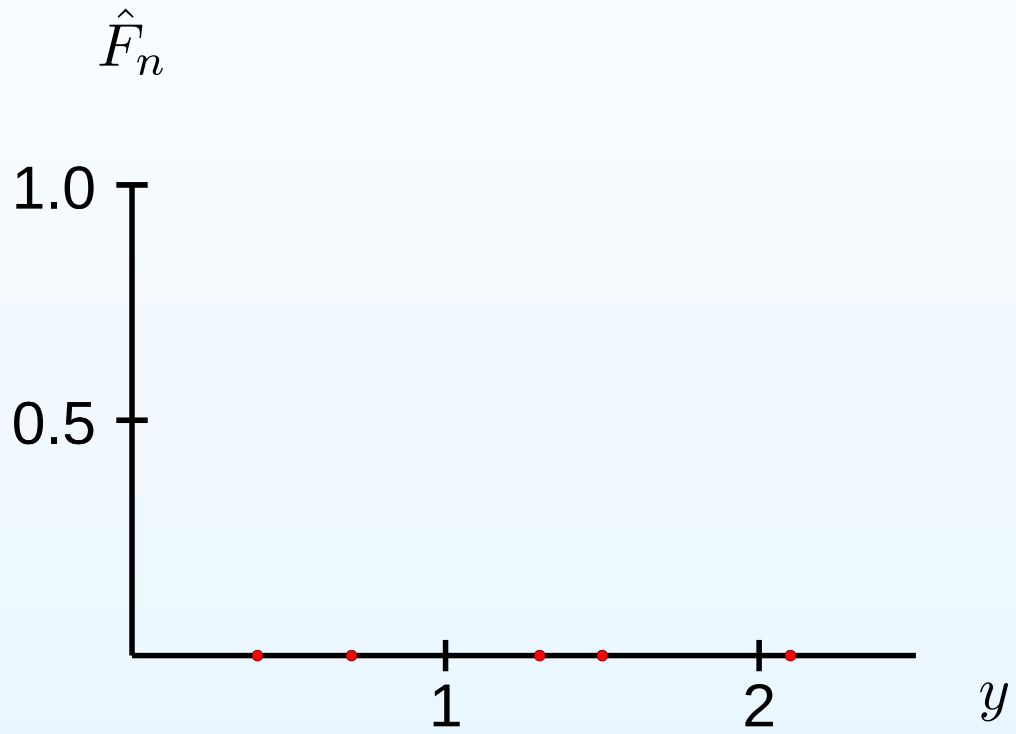
Cumulative Sum Diagram: Example $n = 5$

i	$\Delta_{(i)}$	$Y_{(i)}$
1	1	0.4
2	0	0.7
3	1	1.3
4	1	1.5
5	0	2.0



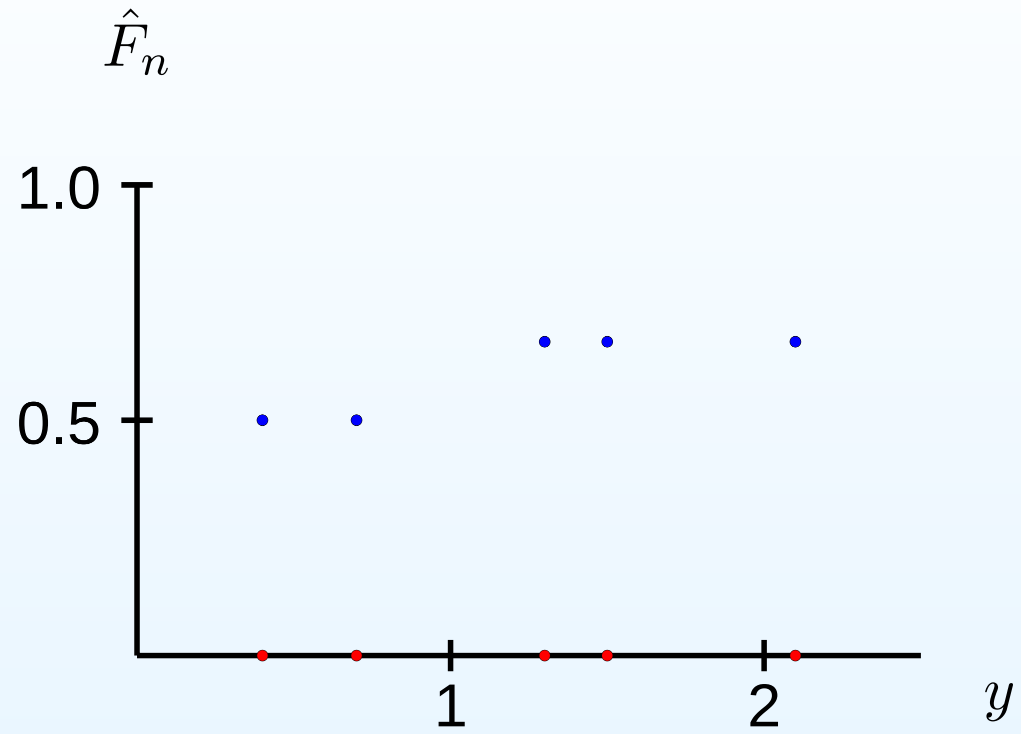
Example continued, $n = 5$

i	$\Delta_{(i)}$	$Y_{(i)}$
1	1	0.4
2	0	0.7
3	1	1.3
4	1	1.5
5	0	2.1



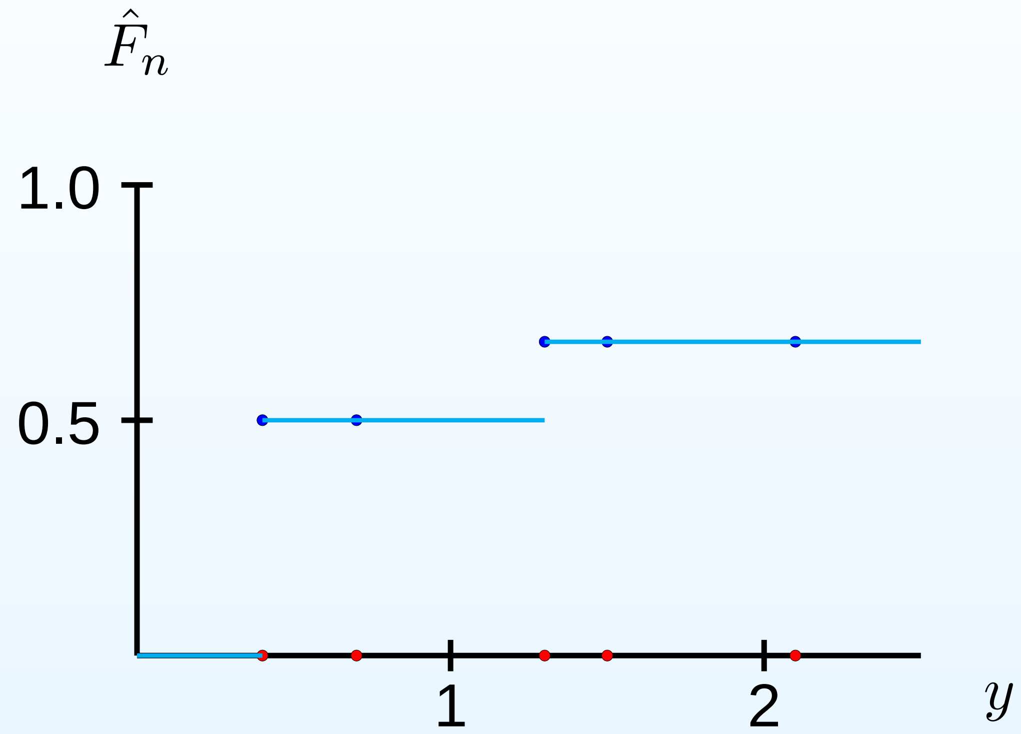
Example continued, $n = 5$

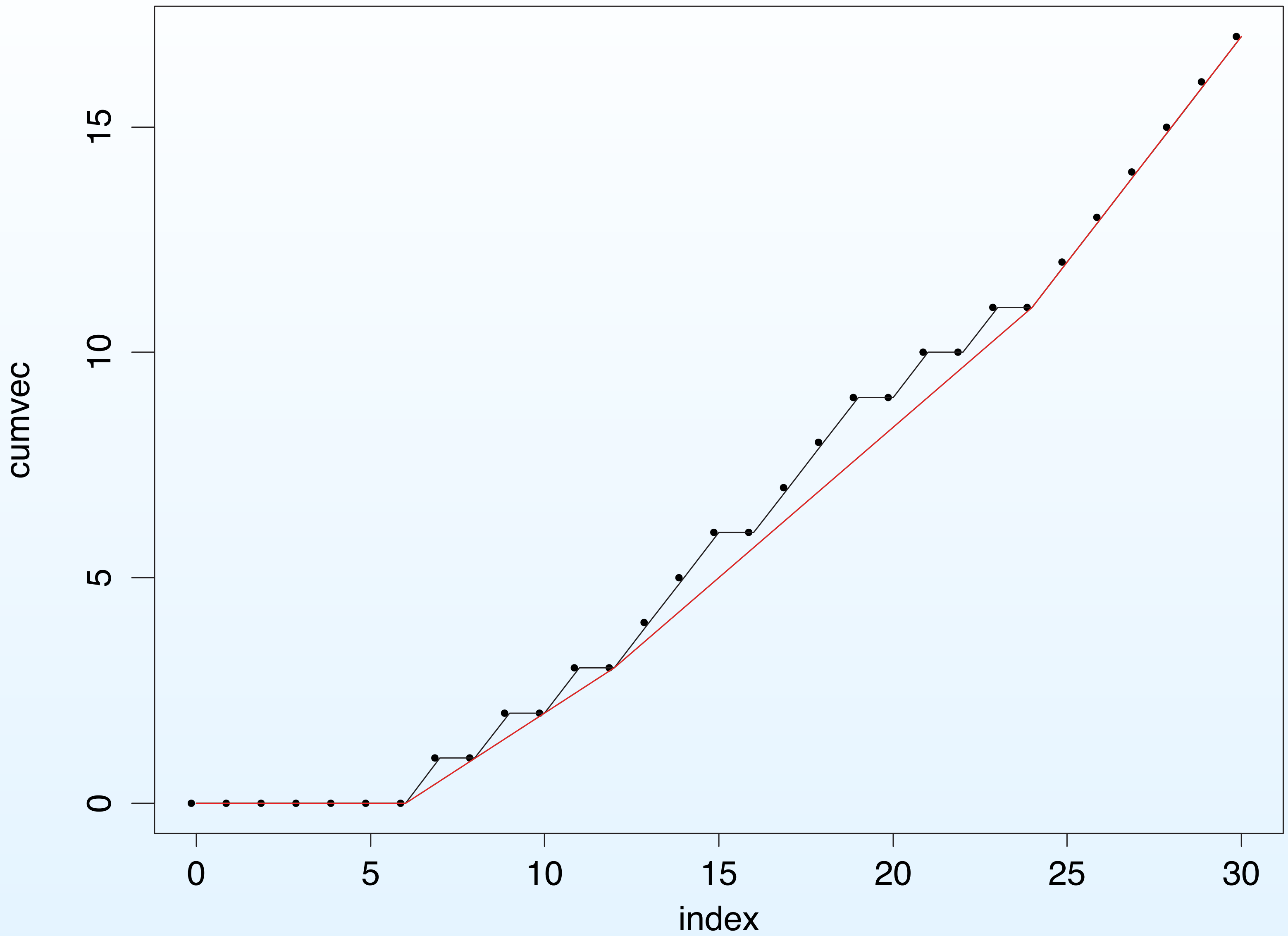
i	$\Delta_{(i)}$	$Y_{(i)}$
1	1	0.4
2	0	0.7
3	1	1.3
4	1	1.5
5	0	2.1

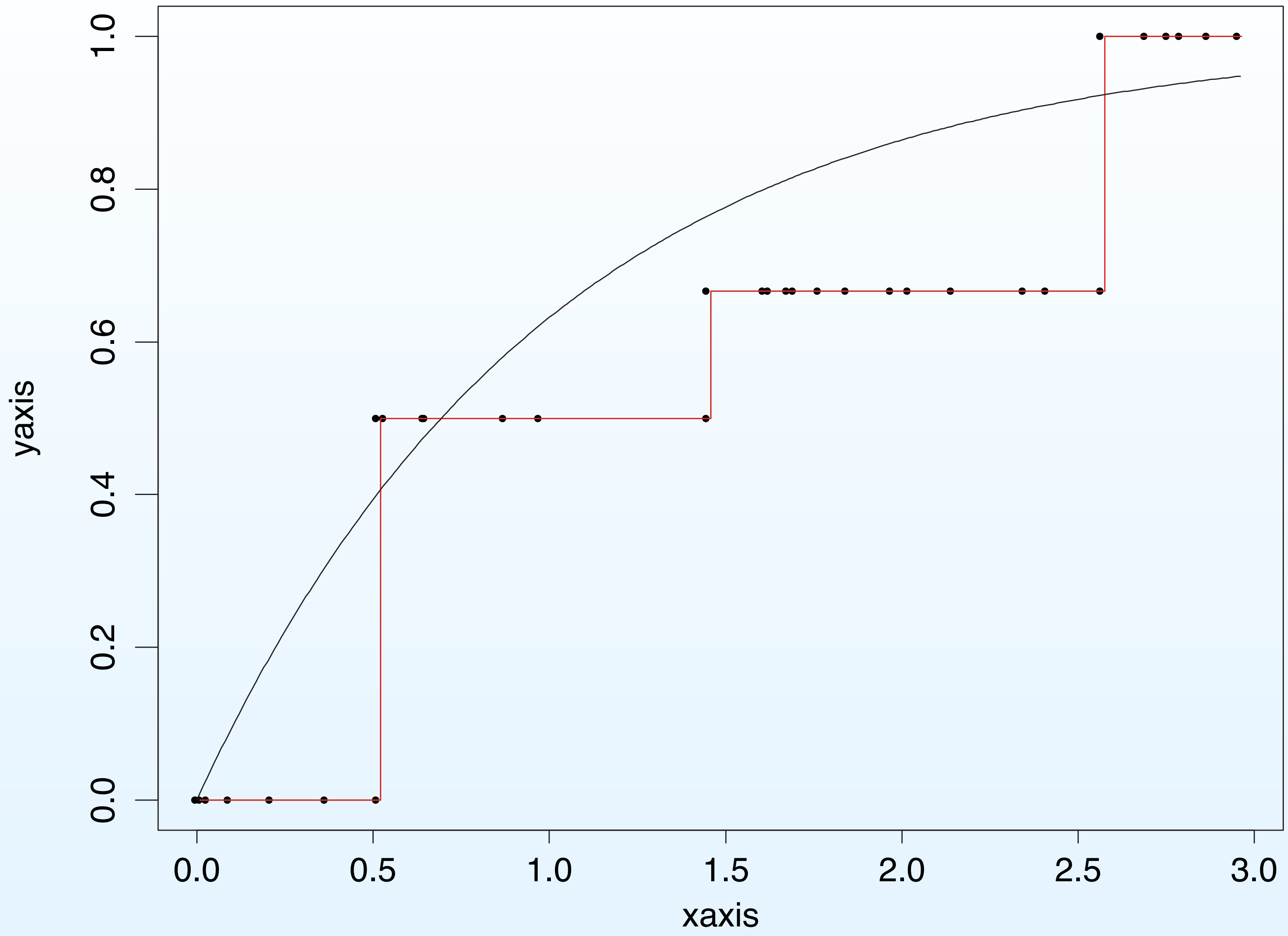


Example continued, $n = 5$

i	$\Delta_{(i)}$	$Y_{(i)}$
1	1	0.4
2	0	0.7
3	1	1.3
4	1	1.5
5	0	2.1




cumvec
index
0
5
10
15
0
5
10
15
20
25
30



1.0
0.8
0.6
0.4
0.2
0.0
yaxis
0.0
0.5
1.0
1.5
2.0
2.5
3.0
xaxis

3. The constrained MLE $\hat{F}_n^0$. Recipe:

- Break $\mathcal{P}$ into $\mathcal{P}_L$ and $\mathcal{P}_R$ where

$$\mathcal{P}_L = \{P_i : \ Y_{(i)} \le t_0\}, \ \ \mathcal{P}_R = \{P_i : \ Y_{(i)} > t_0\} \,.$$

3. The constrained MLE $\hat{F}_n^0$. Recipe:

- Break $\mathcal{P}$ into $\mathcal{P}_L$ and $\mathcal{P}_R$ where

$$\mathcal{P}_L = \{P_i : \ Y_{(i)} \leq t_0\}, \ \ \mathcal{P}_R = \{P_i : \ Y_{(i)} > t_0\}\,.$$

- Form the GCM's of $\mathcal{P}_L$ and $\mathcal{P}_R$, say $\tilde{\mathbb{V}}_n^L$ and $\tilde{\mathbb{V}}_n^R$.

3. The constrained MLE $\hat{F}_n^0$. Recipe:

- Break $\mathcal{P}$ into $\mathcal{P}_L$ and $\mathcal{P}_R$ where

$$\mathcal{P}_L = \{P_i : \ Y_{(i)} \le t_0\}, \ \ \mathcal{P}_R = \{P_i : \ Y_{(i)} > t_0\}.$$

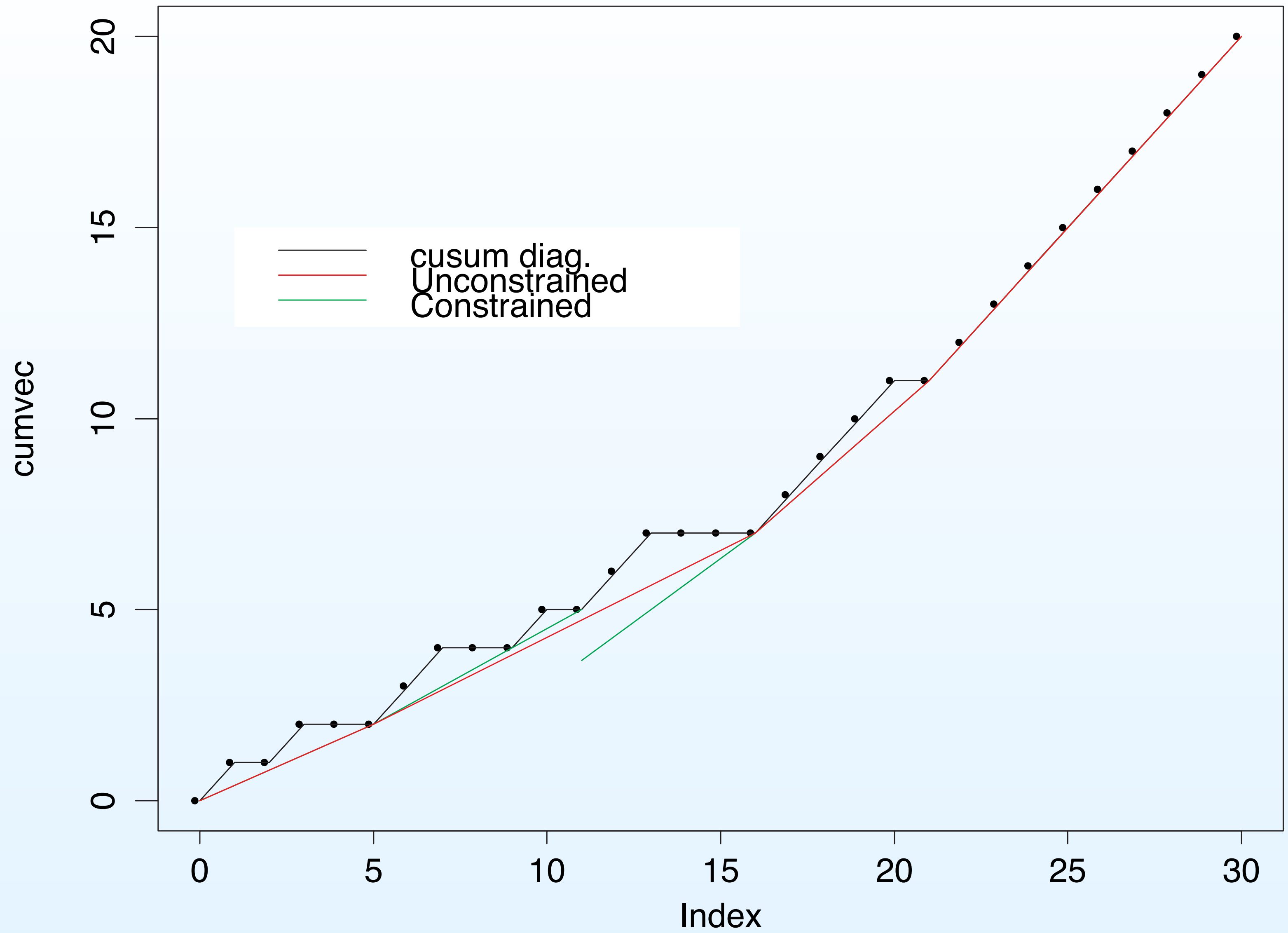
- Form the GCM's of $\mathcal{P}_L$ and $\mathcal{P}_R$, say $\tilde{\mathbb{V}}_n^L$ and $\tilde{\mathbb{V}}_n^R$.
- If the slope of $\tilde{\mathbb{V}}_n^L$ exceeds θ_0, replace it by θ_0; if the slope of $\tilde{\mathbb{V}}_n^R$ drops below θ_0, replace it by θ_0.

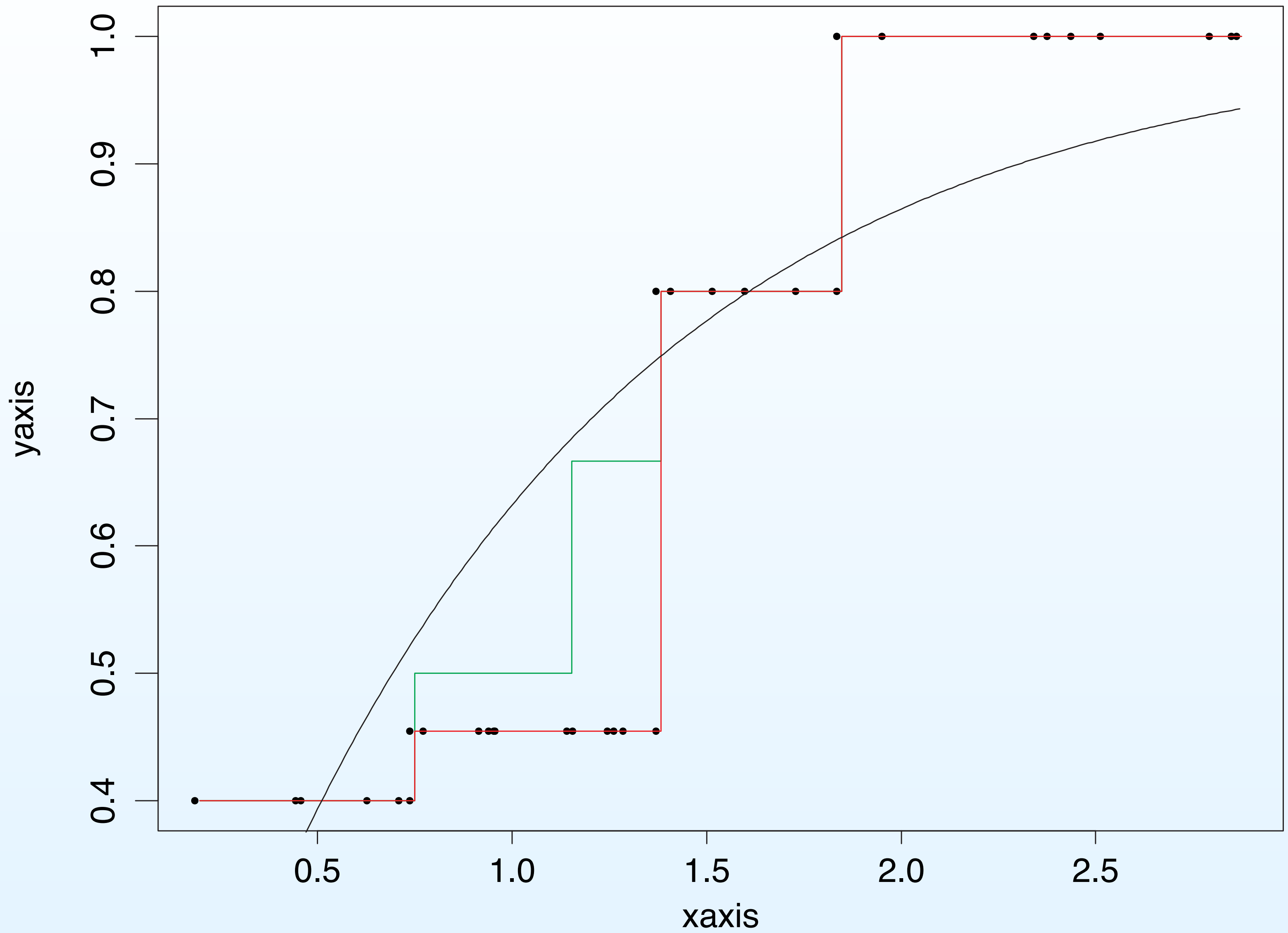
3. The constrained MLE $\hat{F}_n^0$. Recipe:

- Break $\mathcal{P}$ into $\mathcal{P}_L$ and $\mathcal{P}_R$ where

$$\mathcal{P}_L = \{P_i : \; Y_{(i)} \le t_0\}, \;\; \mathcal{P}_R = \{P_i : \; Y_{(i)} > t_0\}\,.$$

- Form the GCM's of $\mathcal{P}_L$ and $\mathcal{P}_R$, say $\tilde{\mathbb{V}}_n^L$ and $\tilde{\mathbb{V}}_n^R$.
- If the slope of $\tilde{\mathbb{V}}_n^L$ exceeds θ_0, replace it by θ_0; if the slope of $\tilde{\mathbb{V}}_n^R$ drops below θ_0, replace it by θ_0.
- The resulting (truncated or constrained) slope process yields the constrained MLE $\hat{F}_n^0$.


cusum diag.
Unconstrained
Constrained
cumvec
Index
0
5
10
15
20
25
30


1.0
0.9
0.8
0.7
0.6
0.5
0.4
yaxis
0.5
1.0
1.5
2.0
2.5
xaxis

4. The likelihood ratio test of $H : F(t_0) = \theta_0$

- Likelihood ratio statistic:

$$\lambda_n = \frac{\sup_F L_n(F)}{\sup_{F:F(t_0)=\theta_0} L_n(F)} = \frac{L_n(\hat{F}_n)}{L_n(\hat{F}_n^0)}.$$

4. The likelihood ratio test of $H : F(t_0) = \theta_0$

- Likelihood ratio statistic:

$$\lambda_n = \frac{\sup_F L_n(F)}{\sup_{F:F(t_0)=\theta_0} L_n(F)} = \frac{L_n(\hat{F}_n)}{L_n(\hat{F}_n^0)}.$$

- How big is “too big”?

4. The likelihood ratio test of $H : F(t_0) = \theta_0$

- Likelihood ratio statistic:

$$\lambda_n = \frac{\sup_F L_n(F)}{\sup_{F:F(t_0)=\theta_0} L_n(F)} = \frac{L_n(\hat{F}_n)}{L_n(\hat{F}_n^0)} .$$

- How big is “too big”?
- When $H : F(t_0) = \theta_0$ holds, does

$$2 \log \lambda_n \to_d \text{something?}$$

4. The likelihood ratio test of $H : F(t_0) = \theta_0$

- Likelihood ratio statistic:

$$\lambda_n = \frac{\sup_F L_n(F)}{\sup_{F:F(t_0)=\theta_0} L_n(F)} = \frac{L_n(\hat{F}_n)}{L_n(\hat{F}_n^0)} .$$

- How big is "too big"?
- When $H : F(t_0) = \theta_0$ holds, does

$$2 \log \lambda_n \rightarrow_d \text{something?}$$

- Answer: Yes! Banerjee and Wellner (2001)

5. How big is “too big”? The limiting Gaussian problem

- Suppose that we observe $\{X(t) : t \in R\}$ where

$$X(t) = F(t) + \sigma W(t)$$

  - $F(t) = \int_{-\infty}^{t} f(s)ds$,
  - f monotone non-decreasing, and
  - W is standard (two-sided) Brownian motion.

5. How big is “too big”? The limiting Gaussian problem

- Suppose that we observe $\{X(t) : t \in R\}$ where

$$X(t) = F(t) + \sigma W(t)$$

  - $F(t) = \int_{-\infty}^{t} f(s)ds$,
  - f monotone non-decreasing, and
  - W is standard (two-sided) Brownian motion.
- Suppose that we want to estimate the monotone function f. Equivalently

$$dX(t) = f(t)dt + \sigma dW(t) \, .$$

5. How big is “too big”? The limiting Gaussian problem

- Suppose that we observe $\{X(t) : t \in R\}$ where

$$X(t) = F(t) + \sigma W(t)$$

  - $F(t) = \int_{-\infty}^{t} f(s)ds$,
  - f monotone non-decreasing, and
  - W is standard (two-sided) Brownian motion.

- Suppose that we want to estimate the monotone function f. Equivalently

$$dX(t) = f(t)dt + \sigma dW(t)\,.$$

- The “canonical monotone function” is a linear one, and we can change σ to 1 by virtue of scaling arguments so the “canonical” version of the problem is as follows:

$$dX(t) = 2tdt + dW(t)\,,$$

- “estimate” $2t$ when $\{X(t) : \ t \in R\}$, is observed. Thus

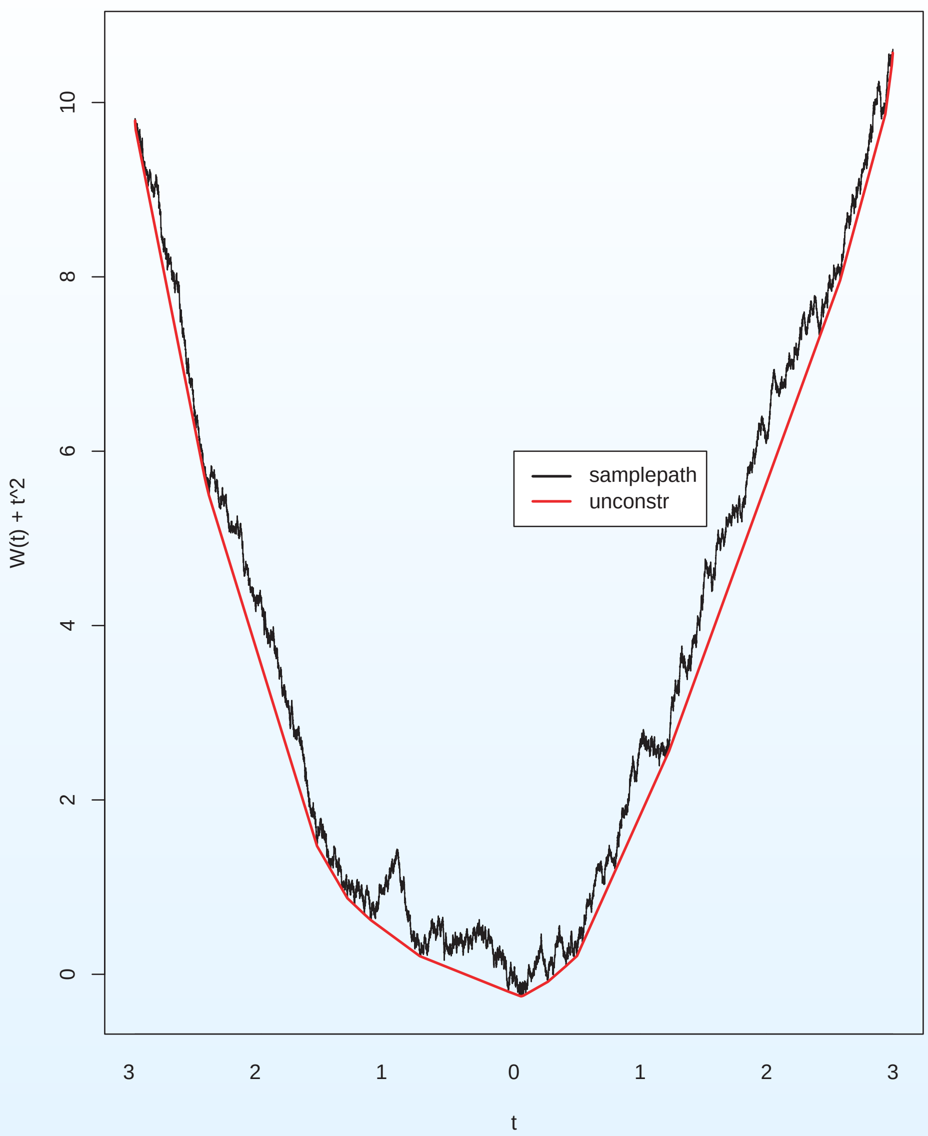
$$X(t) = t^2 + W(t)\,.$$

Unconstrained “Estimator”: Slope of GCM of $X(t)$. Call this process of slopes of the GCM $\mathbb{S}$.

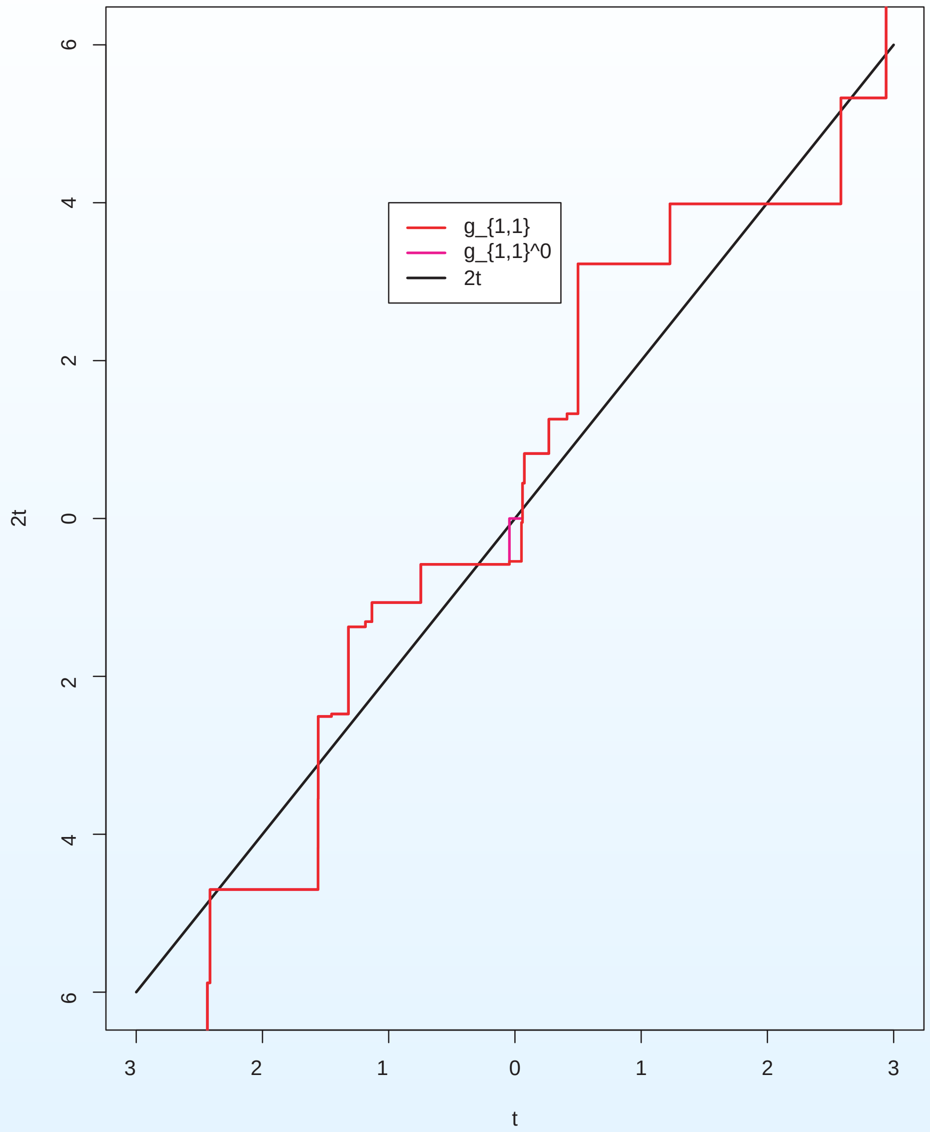
- “estimate” $2t$ when $\{X(t): \ t \in R\}$, is observed. Thus

$$X(t) = t^2 + W(t)\,.$$

Unconstrained “Estimator”: Slope of GCM of $X(t)$. Call this process of slopes of the GCM $\mathbb{S}$.



samplepath
unconstr
W(t) + t^2
t
0
2
4
6
8
10
3
2
1
0
1
2
3

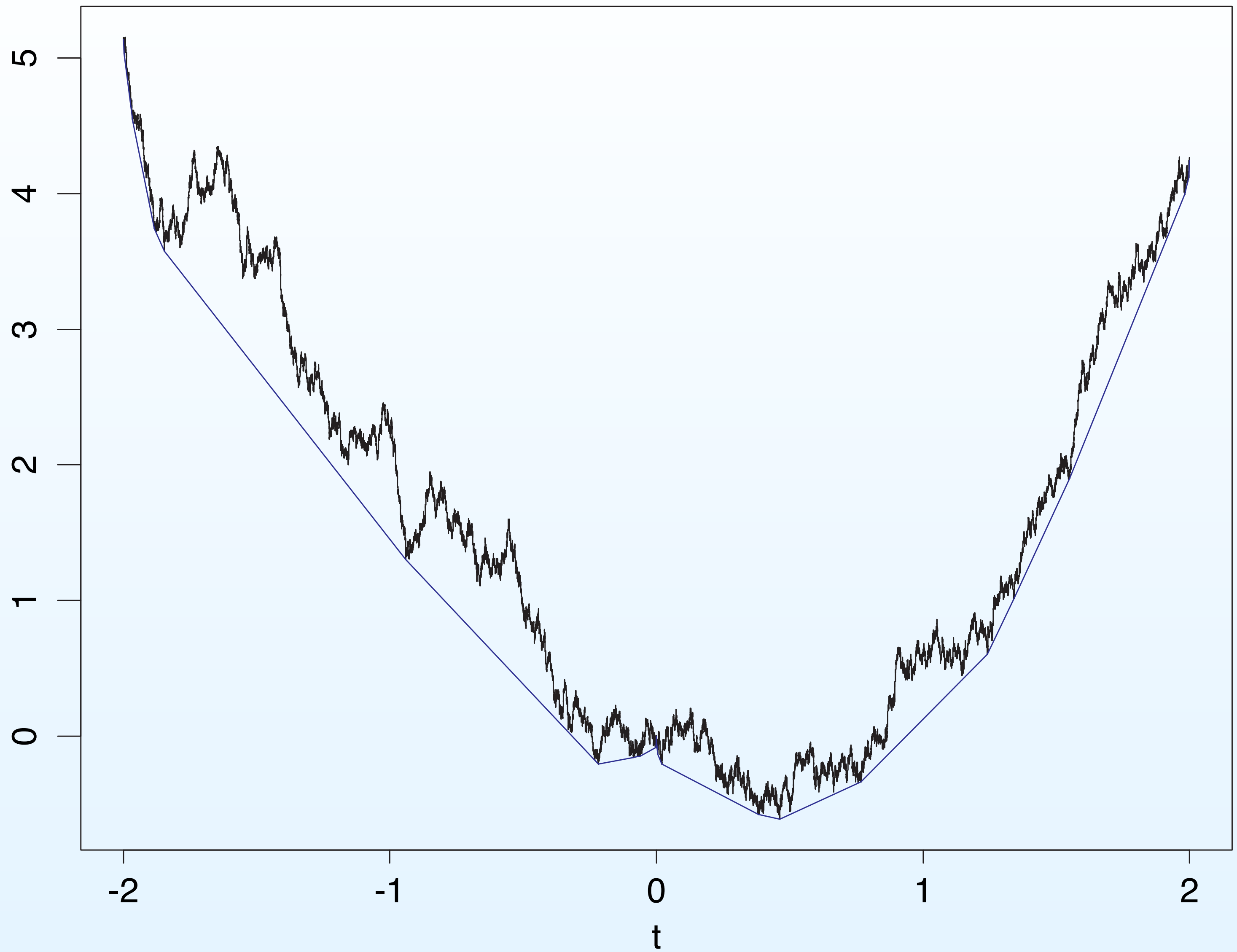

g_{1,1}
g_{1,1}^0
2t
t
2t

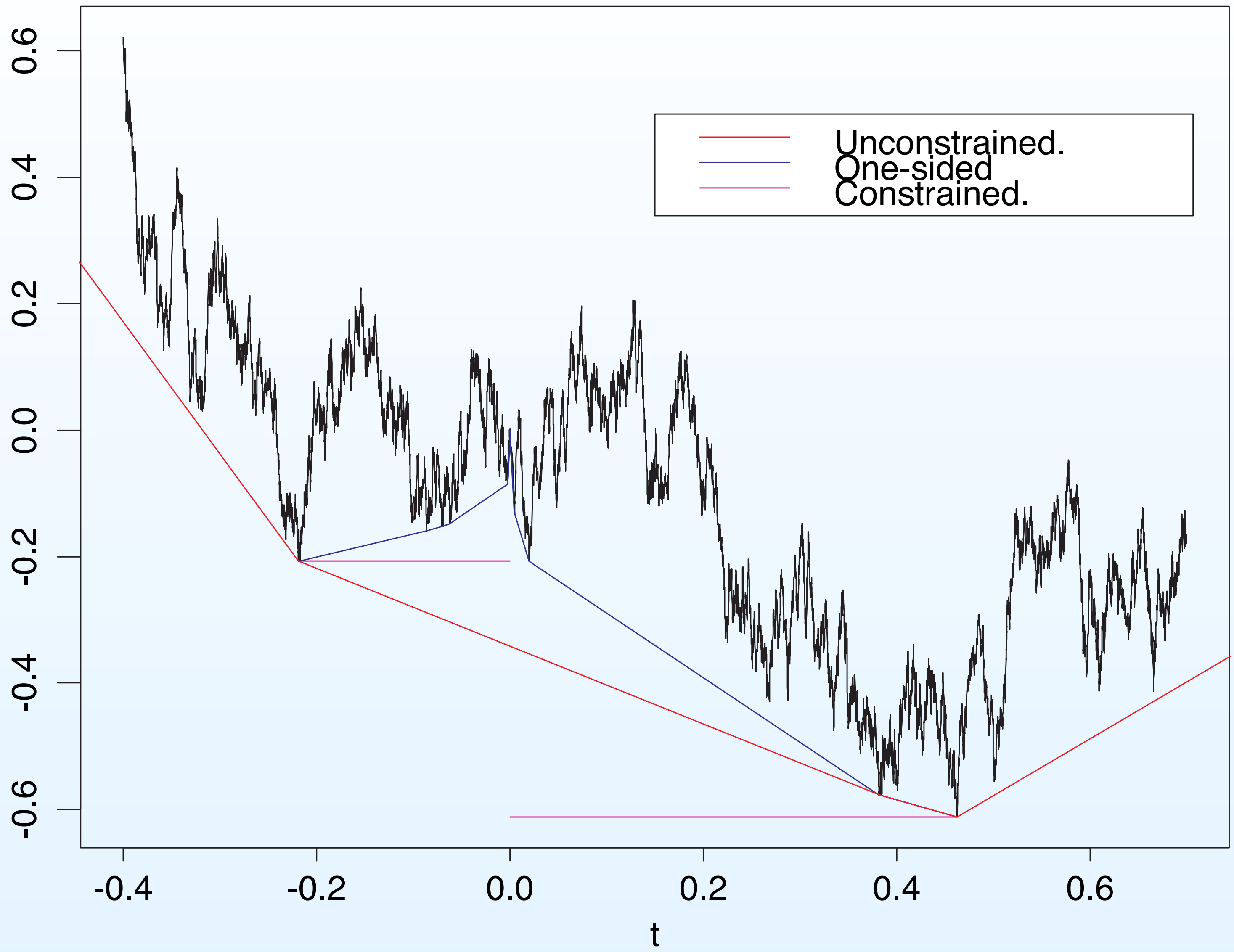
- What is the “canonical constrained problem”?

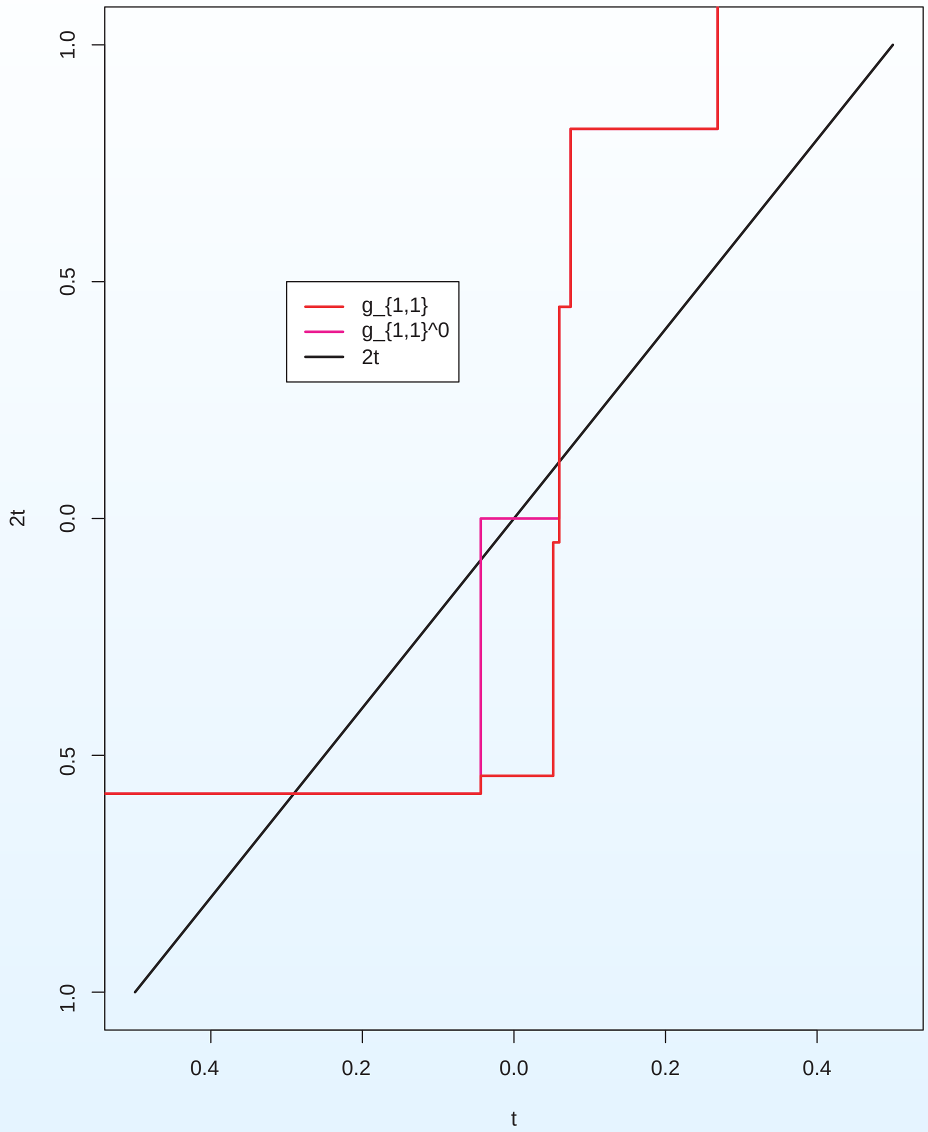
- What is the “canonical constrained problem”?
- Estimate the monontone function $f(t) = 2t$ subject to the constraint that $f(0) = 0$ when $\{X(t) : t \in R\}$ is observed.

- What is the “constrained estimator”?

- What is the “constrained estimator”?
- Recipe:
  - Break $\{X(t) : t \in R\}$ into $X^L \equiv \{X(t) : t < 0\}$ and $X^R \equiv \{X(t) : t \geq 0\}$.
  - Form the GCM’s of X^L and X^R say Y^L and Y^R.
  - If the slope of Y^L exceeds 0, replace it by 0; if the slope of Y^R drops below 0, replace it by 0.
  - The resulting (truncated or constrained) slope process $\mathbb{S}^0$ is the constrained MLE of $f(t) = 2t$ in the Gaussian problem.


5
4
3
2
1
0
-2
-1
0
1
2
t


Unconstrained.
One-sided
Constrained.
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
t


g_{1,1}
g_{1,1}^0
2t
1.0
0.5
0.0
0.5
1.0
0.4
0.2
0.0
0.2
0.4
t

Likelihood ratio test statistic in the Gaussian problem?

- Suppose $\{X(t) : t \in [-c, c]\}$ is given by

$$dX(t) = f(t)dt + dW(t)$$

so

$$X(t) = F(t) + W(t)\,.$$

Likelihood ratio test statistic in the Gaussian problem?

- Suppose $\{X(t) : t \in [-c, c]\}$ is given by

$$dX(t) = f(t)dt + dW(t)$$

so

$$X(t) = F(t) + W(t)\,.$$

- Radon-Nikodym derivative (drifted process relative to zero drift):

$$\frac{dP_f}{dP_0} = \exp\left(\int_{-c}^{c} f\,dX - \frac{1}{2}\int_{-c}^{c} f^2(t)dt\right)\,.$$

Likelihood ratio test statistic in the Gaussian problem?

- Suppose $\{X(t) : t \in [-c, c]\}$ is given by

$$dX(t) = f(t)dt + dW(t)$$

so

$$X(t) = F(t) + W(t)\,.$$

- Radon-Nikodym derivative (drifted process relative to zero drift):

$$\frac{dP_f}{dP_0} = \exp\left(\int_{-c}^{c} f\,dX - \frac{1}{2}\int_{-c}^{c} f^2(t)dt\right)\,.$$

- $\mathcal{F}(c, K) = \{$monotone functions $f : [-c, c] \to \mathbb{R}, \|f\|_c \le K\}$
$\mathcal{F}_0(c, K) = \{f \in \mathcal{F}(c, K) :\ f(0) = 0\}$

- Then

$$
\begin{aligned}
2\log\lambda_c &= 2\log\left(\frac{\sup_{f\in\mathcal{F}(c,K)} dP_f/dP_0}{\sup_{f\in\mathcal{F}_0(c,K)} dP_f/dP_0}\right) = 2\log\left(\frac{dP_{\hat f}/dP_0}{dP_{\hat f_0}/dP_0}\right) \\
&= 2\left\{\int_c^c \hat f_c dX - \frac{1}{2}\int_{-c}^c \hat f_c^2(t)dt \right. \\
&\qquad \left. - \int_c^c \hat f_{c,0} dX + \frac{1}{2}\int_{-c}^c \hat f_{c,0}^2(t)dt\right\} \\
&= 2\int_{-c}^c (\hat f_c - \hat f_{c,0})dX - \int_{-c}^c \{\hat f_c^2(t) - \hat f_{c,0}^2(t)\}dt\,.
\end{aligned}
$$

- Taking the limit as $c \to \infty$ with $K = K_c = 5c$, this yields

$$2 \log \lambda = 2 \int_D (\hat{f} - \hat{f}_0) dX - \int_D \{\hat{f}^2(t) - \hat{f}_0^2(t)\} dt$$

- Taking the limit as $c \to \infty$ with $K = K_c = 5c$, this yields

$$2 \log \lambda = 2 \int_D (\hat{f} - \hat{f}_0) dX - \int_D \{\hat{f}^2(t) - \hat{f}_0^2(t)\} dt$$

- From the characterizations of $\hat{f}$ and $\hat{f}_0$:

$$\int_{\mathbb{R}} (X - \hat{F}) d\hat{f} = 0, \qquad \int_{\mathbb{R}} (X - \hat{F}_0) d\hat{f}_0 = 0\,.$$

- Integration by parts:

$$
\begin{aligned}
\int_{\mathbb{R}} (\hat{f} - \hat{f}_0) dX &= \int_D (\hat{f} - \hat{f}_0) dX = -\int_D X d(\hat{f} - \hat{f}_0) \\
&= -\int_D \hat{F} d\hat{f} + \int_D \hat{F}_0 d\hat{f}_0 \\
&= \int_D \hat{f} d\hat{F} - \int_D \hat{f}_0 d\hat{F}_0 \\
&= \int_D \{\hat{f}^2(t) - \hat{f}_0^2(t)\} dt \,.
\end{aligned}
$$

- Integration by parts:

$$\begin{aligned}\int_{\mathbb{R}} (\hat{f} - \hat{f}_0) dX &= \int_D (\hat{f} - \hat{f}_0) dX = -\int_D X d(\hat{f} - \hat{f}_0) \\ &= -\int_D \hat{F} d\hat{f} + \int_D \hat{F}_0 d\hat{f}_0 \\ &= \int_D \hat{f} d\hat{F} - \int_D \hat{f}_0 d\hat{F}_0 \\ &= \int_D \{\hat{f}^2(t) - \hat{f}_0^2(t)\} dt\,.\end{aligned}$$

- Likelihood ratio statistic becomes:

$$2\log\lambda = \int_D \{\hat{f}^2(t) - \hat{f}_0^2(t)\} dt\,.$$

6. Limit distribution, LR statistic under H

- Limit distributions for $\hat{F}_n$ and $\hat{F}_n^0$. Set

$$\mathbb{G}_n^{loc}(t,h) = n^{1/3}(\mathbb{G}_n(t+n^{-1/3}h) - \mathbb{G}_n(t))$$

$$\mathbb{V}_n^{loc}(t,h)$$

$$= n^{1/3}\left\{n^{1/3}(\mathbb{V}_n(t+n^{-1/3}h) - \mathbb{V}_n(t)) - \mathbb{G}_n^{loc}(t,h)F(t)\right\}.$$

6. Limit distribution, LR statistic under H

- Limit distributions for $\hat{F}_n$ and $\hat{F}_n^0$. Set

$$\mathbb{G}_n^{loc}(t,h) = n^{1/3}(\mathbb{G}_n(t+n^{-1/3}h) - \mathbb{G}_n(t))$$
$$\mathbb{V}_n^{loc}(t,h)$$
$$= n^{1/3}\left\{n^{1/3}(\mathbb{V}_n(t+n^{-1/3}h) - \mathbb{V}_n(t)) - \mathbb{G}_n^{loc}(t,h)F(t)\right\} .$$

- **Theorem 1.** If $g(t_0) = G'(t_0)$ and $f(t_0) = F'(t_0)$ exist, then:
  A. $\mathbb{G}_n^{loc}(t_0,h) \to_p g(t_0)h$.
  B. $\mathbb{V}_n^{loc}(t_0,h) \Rightarrow aW(h) + bh^2$ where
  $a = \sqrt{F(t_0)(1-F(t_0))g(t_0)}$, $b = f(t_0)g(t_0)/2$, and W is a two-sided Brownian motion starting from 0.

- Now define

$$\mathbb{Z}_n(h) = n^{1/3}(\hat{F}_n(t_0 + hn^{-1/3}) - F(t_0)) \,,$$
$$\mathbb{Z}_n^0(h) = n^{1/3}(\hat{F}_n^0(t_0 + hn^{-1/3}) - F(t_0)) \,.$$

- Now define

$$\mathbb{Z}_n(h) = n^{1/3}(\hat{F}_n(t_0 + hn^{-1/3}) - F(t_0)) \, ,$$
$$\mathbb{Z}_n^0(h) = n^{1/3}(\hat{F}_n^0(t_0 + hn^{-1/3}) - F(t_0)) \, .$$

- **Theorem 2.** If the hypotheses of Theorem 1 hold with $f(t_0) > 0$, $g(t_0) > 0$, and $F(t_0) = \theta_0$, then

$$(\mathbb{Z}_n(h), \mathbb{Z}_n^0(h)) \Rightarrow (\mathbb{S}_{a,b}(h), \mathbb{S}_{a,b}^0(h))/g(t_0)$$

where $\mathbb{S}_{a,b}$ and $\mathbb{S}_{a,b}^0$ are the constrained and unconstrained slope processes corresponding to $X_{a,b}(h) = aW(h) + bh^2$.

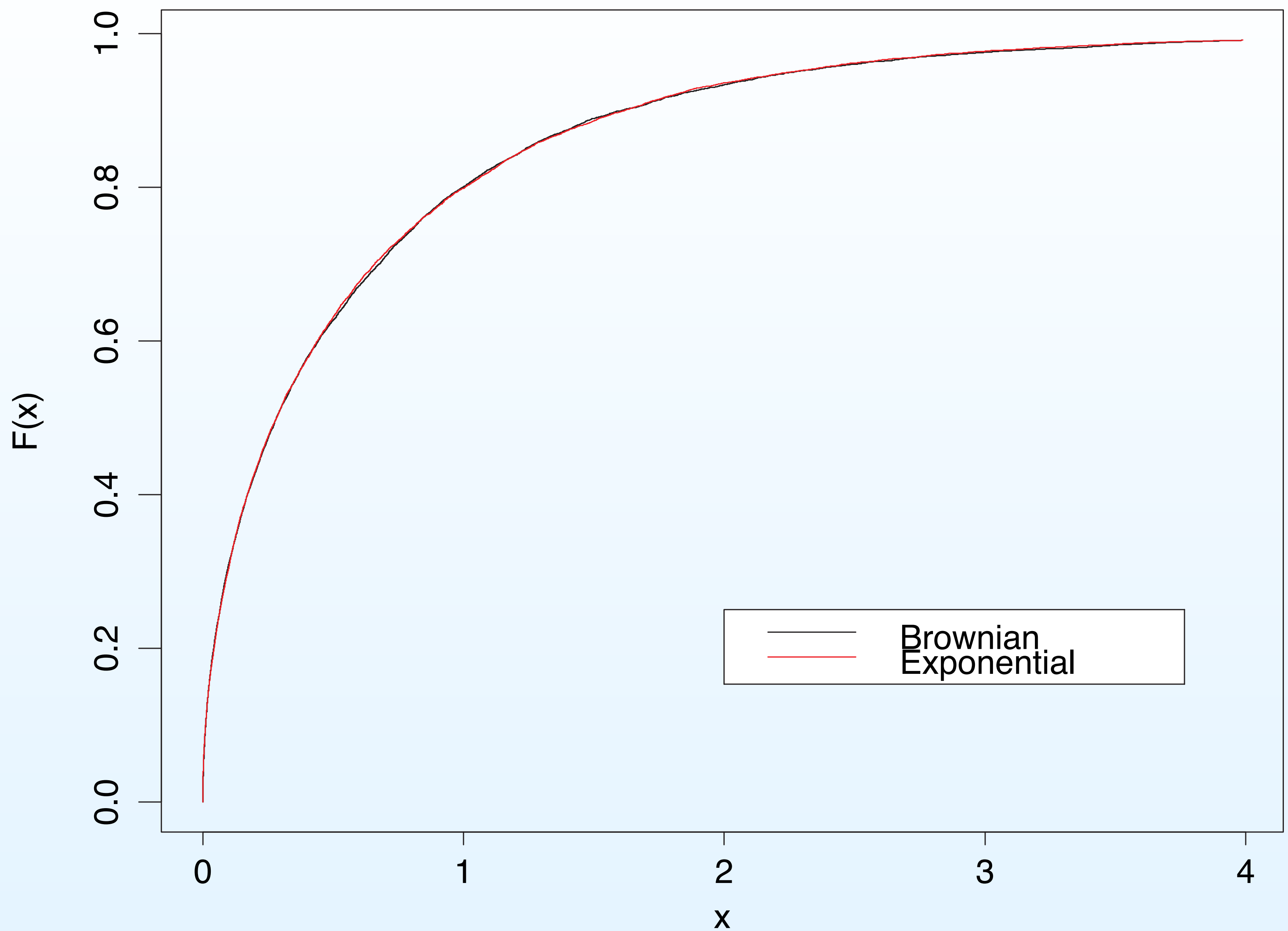
- Limit distribution for $2 \log \lambda_n$

- Limit distribution for $2\log\lambda_n$
- **Theorem 3.** (Banerjee and Wellner, 2001). Suppose that F and G have densities f and g which are strictly positive and continuous in a neighborhood in a neighborhood of t_0. Suppose that $F(t_0) = \theta_0$. Then

$$2\log\lambda_n \quad \to_d \quad \frac{1}{g(t_0)a^2}\int ((\mathbb{S}_{a,b}(z))^2 - (\mathbb{S}^0_{a,b}(z))^2)\, dz$$

$$\stackrel{d}{=} \quad \int \{(\mathbb{S}(z))^2 - (\mathbb{S}^0(z))^2\} dz \equiv \mathbb{D}\,,$$

and the distribution of $\mathbb{D}$ is universal (free of parameters).


1.0
0.8
0.6
0.4
0.2
0.0
F(x)
0
1
2
3
4
x
Brownian
Exponential

7. Confidence intervals for $F(t_0)$

- Wald-type intervals:

$$
\begin{aligned}
\mathbb{Z}_n(0) &= n^{1/3}(\hat{F}_n(t_0) - F(t_0)) \to_d \mathbb{S}_{a,b}(0)/g(t_0) \\
&\stackrel{d}{=} \left\{ \frac{F(t_0)(1-F(t_0))f(t_0)}{2g(t_0)} \right\}^{1/3} \mathbb{S}(0) \\
&\equiv C(F,f,g)\,\mathbb{S}(0)
\end{aligned}
$$

where $\mathbb{S}(0) \stackrel{d}{=} 2\,\mathbb{Z} \equiv 2\text{argmin}(W(h) + h^2)$, $\mathbb{S}(0) \equiv \mathbb{S}_{1,1}(0)$.

7. Confidence intervals for $F(t_0)$

- Wald-type intervals:

$$\begin{aligned}\mathbb{Z}_n(0) &= n^{1/3}(\hat{F}_n(t_0) - F(t_0)) \rightarrow_d \mathbb{S}_{a,b}(0)/g(t_0) \\ &\stackrel{d}{=} \left\{ \frac{F(t_0)(1-F(t_0))f(t_0)}{2g(t_0)} \right\}^{1/3} \mathbb{S}(0) \\ &\equiv C(F,f,g)\,\mathbb{S}(0)\end{aligned}$$

where $\mathbb{S}(0) \stackrel{d}{=} 2\,\mathbb{Z} \equiv$ 2argmin$(W(h)+h^2)$, $\mathbb{S}(0) \equiv \mathbb{S}_{1,1}(0)$.

- Wald - interval:

$$\hat{F}_n(t_0) \pm n^{-1/3} C(\hat{F}_n, \hat{f}_n, \hat{g}_n)\, t_\alpha$$

where $\hat{f}_n$ and $\hat{g}_n$ are estimates of f and g (at t_0), and $t_{\alpha/2}$ satisfies

$$P(2\mathbb{Z} > t_{\alpha/2}) = \alpha/2\,.$$

- Problem: this involves smoothing to get estimators $\hat{f}_n$ and $\hat{g}_n$!

- Confidence intervals from the LR test

- Confidence intervals from the LR test
- Invert the test:

$$\{\theta : 2\log\lambda_n(\theta) \leq d_\alpha\}.$$

where $P(\mathbb{D} \leq d_\alpha) = 1 - \alpha$

- Confidence intervals from the LR test
- Invert the test:

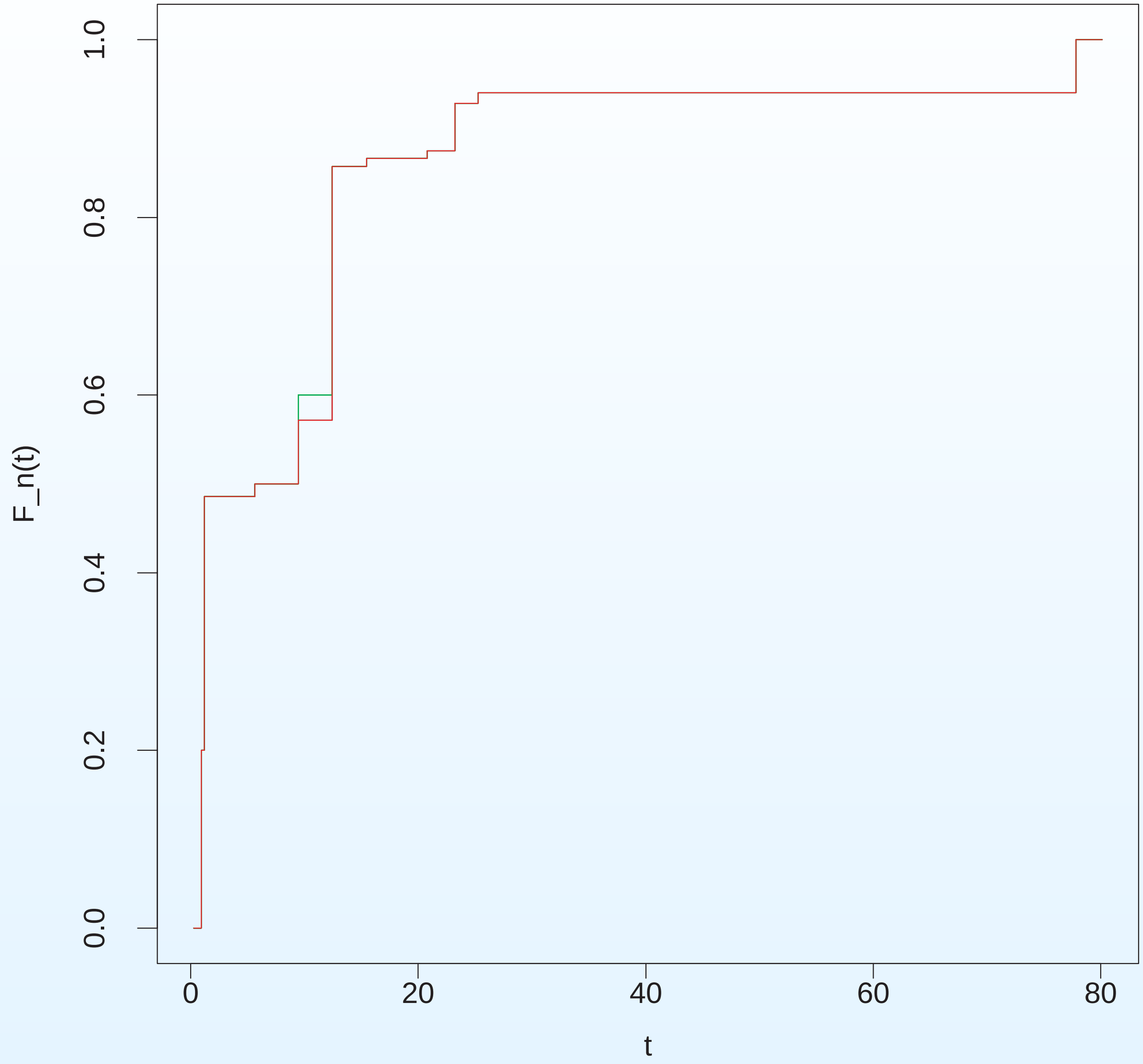
$$\{\theta : 2\log\lambda_n(\theta) \leq d_\alpha\}\,.$$

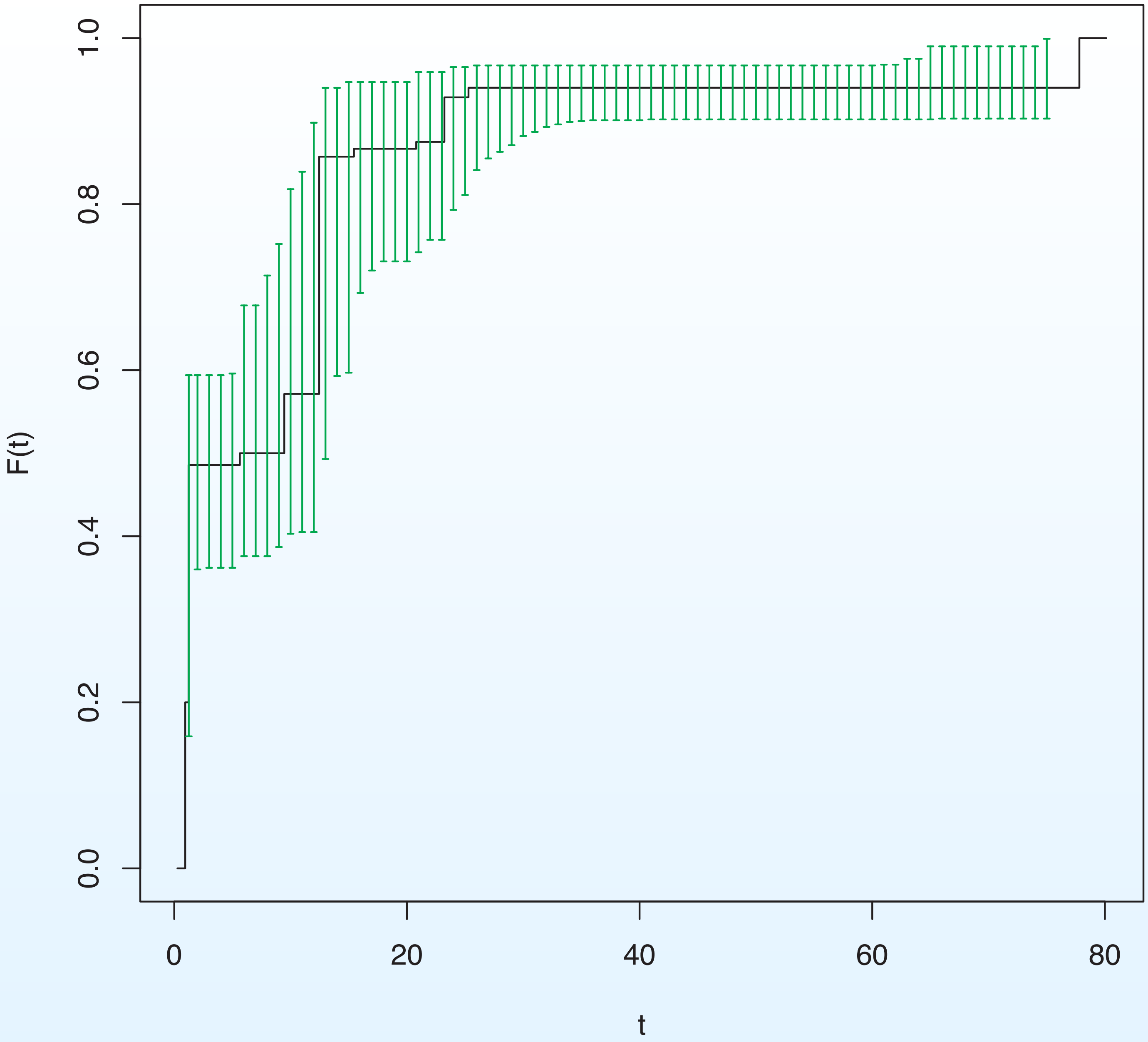
  where $P(\mathbb{D} \leq d_\alpha) = 1 - \alpha$
- Advantage: no smoothing needed!

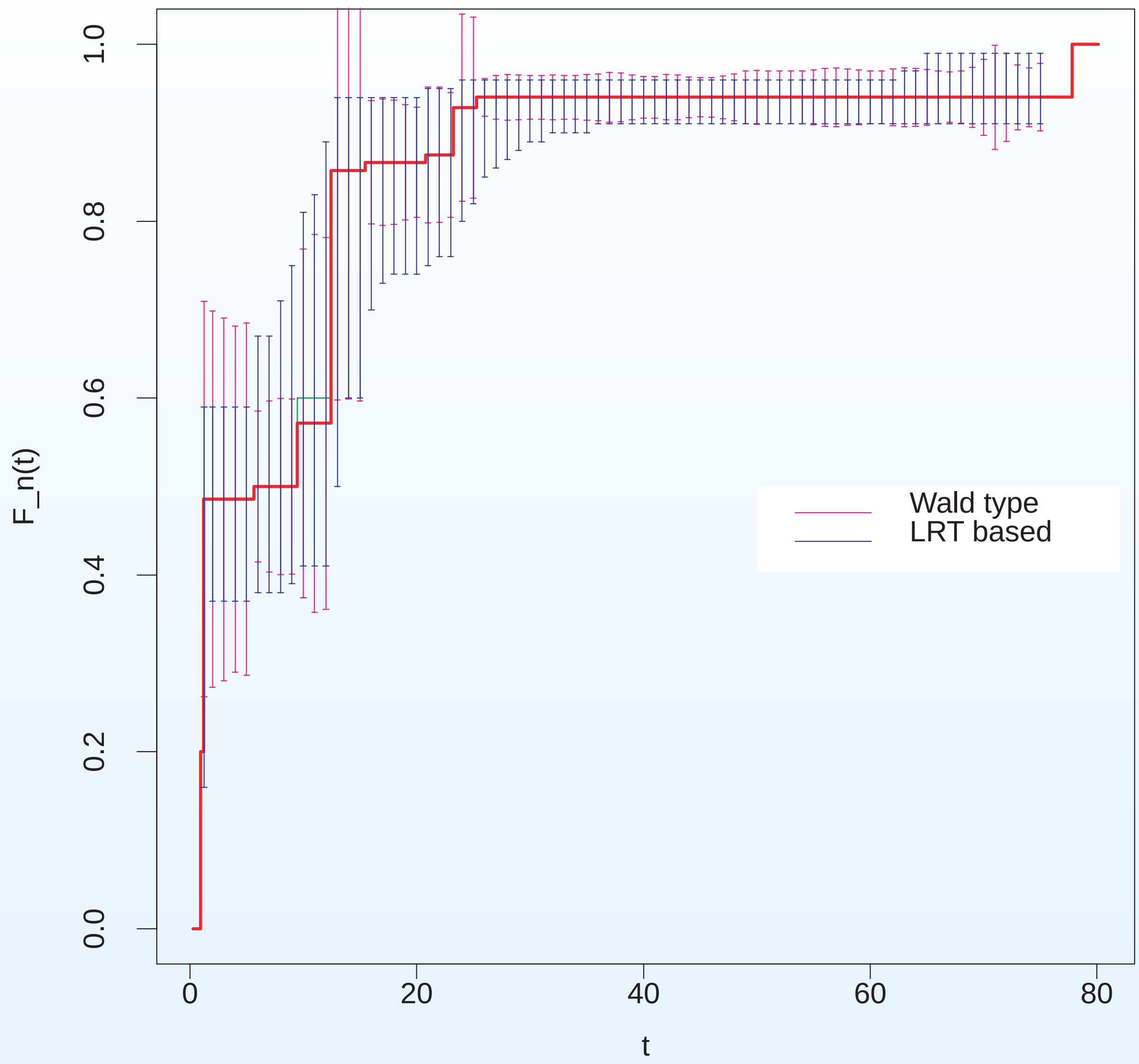
- Confidence intervals from the LR test
- Invert the test:

$$\{\theta : 2\log\lambda_n(\theta) \le d_\alpha\}\,.$$

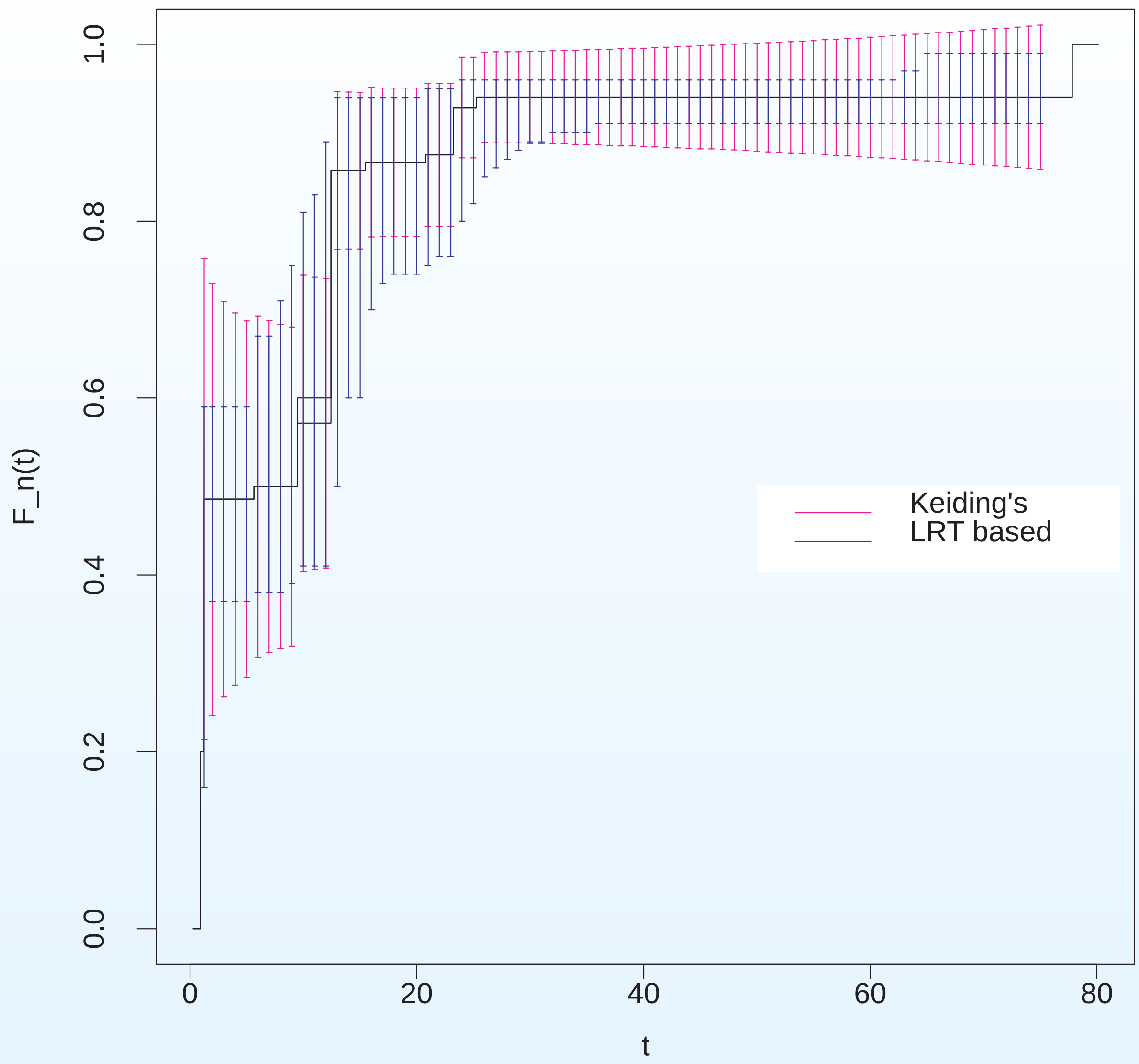
  where $P(\mathbb{D} \le d_\alpha) = 1 - \alpha$
- Advantage: no smoothing needed!
- Tradeoff: need to compute constrained estimator(s) $\hat{F}_n^0$ of F and $\lambda_n(\theta)$ for many different values of the constraint θ.


F_n(t)
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
t


F(t)
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
t



F_n(t)
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
t
Wald type
LRT based


F_n(t)
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
t
Keiding's
LRT based

8. Further problems

- Can we find the distribution of $\mathbb{D}$ analytically?

8. Further problems

- Can we find the distribution of $\mathbb{D}$ analytically?
- Does the same limit $\mathbb{D}$ arise as the limit distribution for the likelihood ratio test for a large class of such problems involving monotone functions? (Yes – Banerjee (2005))

8. Further problems

- Can we find the distribution of $\mathbb{D}$ analytically?
- Does the same limit $\mathbb{D}$ arise as the limit distribution for the likelihood ratio test for a large class of such problems involving monotone functions? (Yes – Banerjee (2005))
- Power? What is the appropriate contiguity theory? What is the limit distribution of the likelihood ratio statistic under local alternatives? See Banerjee and Wellner (2001), (2005b).

8. Further problems

- Can we find the distribution of $\mathbb{D}$ analytically?
- Does the same limit $\mathbb{D}$ arise as the limit distribution for the likelihood ratio test for a large class of such problems involving monotone functions? (Yes – Banerjee (2005))
- Power? What is the appropriate contiguity theory? What is the limit distribution of the likelihood ratio statistic under local alternatives? See Banerjee and Wellner (2001), (2005b).
- What happens if we constrain at $k > 1$ points? (Independence!)

8. Further problems

- Can we find the distribution of $\mathbb{D}$ analytically?
- Does the same limit $\mathbb{D}$ arise as the limit distribution for the likelihood ratio test for a large class of such problems involving monotone functions? (Yes – Banerjee (2005))
- Power? What is the appropriate contiguity theory? What is the limit distribution of the likelihood ratio statistic under local alternatives? See Banerjee and Wellner (2001), (2005b).
- What happens if we constrain at $k > 1$ points? (Independence!)
- Confidence bands for the whole monotone function F?

8. Further problems

- Can we find the distribution of $\mathbb{D}$ analytically?
- Does the same limit $\mathbb{D}$ arise as the limit distribution for the likelihood ratio test for a large class of such problems involving monotone functions? (Yes – Banerjee (2005))
- Power? What is the appropriate contiguity theory? What is the limit distribution of the likelihood ratio statistic under local alternatives? See Banerjee and Wellner (2001), (2005b).
- What happens if we constrain at $k > 1$ points? (Independence!)
- Confidence bands for the whole monotone function F?
- Confidence intervals (and bands?) for estimating a concave distribution function F?