Introduction to Sequential Change-Point Problems

Yajun Mei

ymei@fhcrc.org

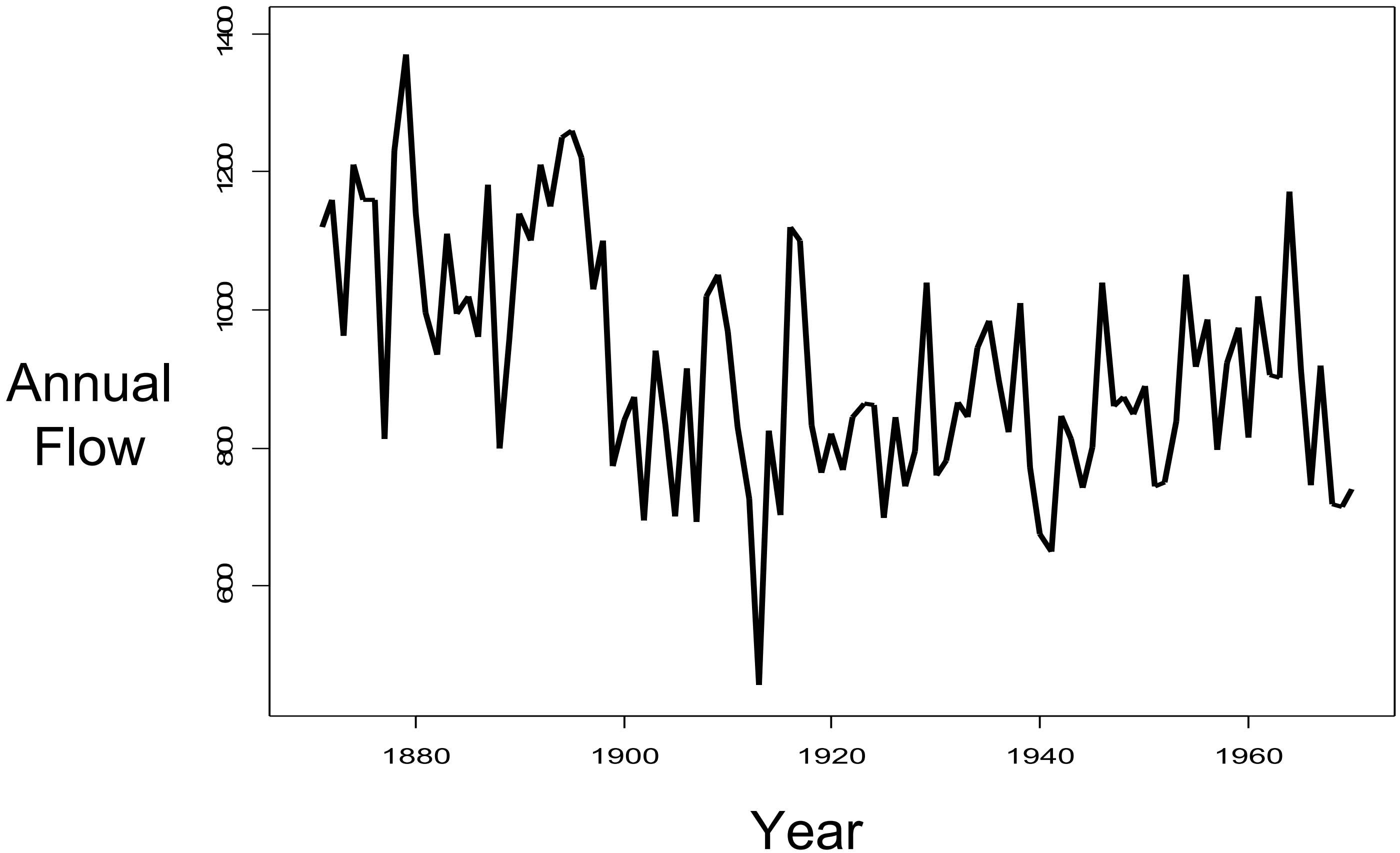
SAMSI
September 29, 2005

Outline

- Example: Nile River
- Sequential Change-Point Problems
  - Minimax Formulation
  - Bayesian Formulation
- Generalizations

Nile River (1871-1970)

Cobb (*Biometrika*, 1978)



Applications

- Quality / Process control
- Epidemiology
- Signal processing
- Finance
- Surveillance / Security
- Others ?

Change-Point Problem Formulation

$$\underbrace{X_1, X_2, \cdots, X_{\nu-1}}_{\text{i.i.d. } f} \longrightarrow \underbrace{X_\nu, X_{\nu+1}, \cdots}_{\text{i.i.d. } g}$$

- ***Goal:*** **Raise an alarm as soon as a change occurs**
- **A procedure is defined as a stopping time** T

($T < \infty$ $\longrightarrow$ **declare a change has occurred)**

Minimax Formulation

- **Detection Delay:**

$$D(T) = \sup_{1 \leq \nu < \infty} \mathbf{E}[\mathbf{T} - \nu | \mathbf{T} \geq \nu]$$

- **False Alarm Rate:**
  - Can P(ever raise a false alarm) $\leq$ 5%?

    No! Lorden (*Ann. Math. Stat.*, 1971) showed:

    D(T) is finite $\rightarrow$ P(raise a false alarm) = 1
  - Usually measured by $1/\mathbf{E_f}[\mathbf{T}]$, where $\mathbf{E_f}[\mathbf{T}]$ is Mean Time until a False Alarm (MTFA)

An Instructive Example

Before Change (B. C.), X_i's are i.i.d. N(0, 1)
After Disorder (A. D.), X_i's are i.i.d. N(1, 1)

<u>Problem:</u> Minimize detection delay D(T) subject to MTFA $\geq \gamma$ (e.g., =100)

- **CUSUM procedure (Page (1954))**

  T_{CM} = first n such that $W_n \geq 2.85$, where

$$W_n = \max_{1 \leq k \leq n} \sum_{i=k}^{n} (X_i - 0.5) \quad [= \max\{0, W_{n-1}\} + (X_n - 0.5)]$$

- **T_{CM} is (nearly) optimal: $D(T_{CM}) \approx 6.1$**

Page's CUSUM Procedures

$$\underbrace{X_1, X_2, \cdots, X_{\nu-1}}_{\text{i.i.d. } f} \longrightarrow \underbrace{X_\nu, X_{\nu+1}, \cdots}_{\text{i.i.d. } g}$$

- **Given $X_1, \cdots, X_n$, the log-likelihood ratio of**

  $H_0 : \nu = \infty$ vs. $H_{1,k} : \nu = k$ **is** $\sum_{i=k}^{n} \log \frac{g(X_i)}{f(X_i)}$

- **CUSUM statistics is Maximum Likelihood Ratio**

$$W_n = \max_{1 \le k \le n} \sum_{i=k}^{n} \log \frac{g(X_i)}{f(X_i)}$$

- **Page's CUSUM = first n such that** $W_n \ge a$

- **(Asymptotic) optimality: (Lorden 1971, Moustakides, 1986)**

Bayesian Formulation

- **The change-point ν is a random variable with a known prior distribution**

- **Problem:** **Minimize** $\mathsf{P}_{\mathbf{f}}(\mathbf{T} < \nu) + \mathbf{c}\mathsf{E}(\mathbf{T} - \nu)^{+}$

  where c >0 is a pre-observation cost of delay.

- **Solution:** **If the prior for ν is geometric(p), Bayes rule is**

  $T_{p,\,c}$ = first n such that

$$\mathsf{P}(\nu \leq \mathbf{n} \mid \mathbf{X_1}, \cdots, \mathbf{X_n}) \geq \delta_{\mathbf{p},\mathbf{c}}$$

Shiryeyev-Roberts Procedures

$$\underbrace{X_1, X_2, \cdots, X_{\nu-1}}_{\text{i.i.d.} \quad f} \rightarrow \underbrace{X_\nu, X_{\nu+1}, \cdots}_{\text{i.i.d.} \quad g}$$

- **Shiryayev-Roberts = first n such that** $R_n \geq A$

$$R_n = \sum_{k=1}^{n} \prod_{i=k}^{n} \frac{g(X_i)}{f(X_i)}$$

- **It is the limit of Bayes rules as p $\rightarrow$ 0**
- **Minimax optimality (Pollak, 1985)**

Page's CUSUM & Shiryayev-Roberts

$$\underbrace{X_1, X_2, \cdots, X_{\nu-1}}_{\text{i.i.d.} \quad f} \longrightarrow \underbrace{X_\nu, X_{\nu+1}, \cdots}_{\text{i.i.d.} \quad g}$$

- **CUSUM statistic:**

$$W_n = \max_{1 \le k \le n} \prod_{i=k}^{n} \frac{g(X_i)}{f(X_i)} = \max(W_{n-1}, 1)\frac{g(X_n)}{f(X_n)}$$

- **Shiryayev-Roberts statistic:**

$$R_n = \sum_{k=1}^{n} \prod_{i=k}^{n} \frac{g(X_i)}{f(X_i)} = (R_{n-1} + 1)\frac{g(X_n)}{f(X_n)}$$

- **Their performances are similar under minimax criteria**

Generalizations

- **Pre-change and/or post-change distributions involve unknown parameters** (Lorden 1971; Pollak 1987; Pollak & Siegmund 1991; Lai 1995; Yakir 1998; Gordon & Pollak 1997;Baron 2000; Mei 2003; Krieger, Pollak & Yakir 2003;………)
- **Dependent observations; Hidden Markov** (Lai 1998; Fuh 2003, 2004)
- **Wiener process; Poisson process; Compound Poisson process** (Shiryeyev 1978; Gal'chuk & Rozovkii 1971; Gapeev 2005)
- **Exponential penalty for delay** (Poor 1998; Beibel 2000)
- **Joint detect & isolate changes** (Nikiforov 1995; Lai 2000)
- **Decentralized systems** (Veeravalli 2001; Mei 2005)