Observing process which may change abruptly at unknown, random time.
Want to detect change ASAP after it occurs, while minimizing false alarms. Procedure good for detecting small persistent changes.

Applications
1) Industrial process control: has manufacturing process gone out of control, too many defectives?
2) Lab quality control
   Controls (known values) included in each batch of samples. Equipment problem may change level or variance of measurements. Detect by observing values for controls across several batches.
3) Medical quality control: Is a surgical unit starting to have excess fatalities?
4) Health surveillance: Is rare disease becoming epidemic?
5) Psychology learning experiment: Is there a trial at which subject suddenly gets insights, behavior changes?

Necessarily sequential
Cannot be done under FSS.
Stat. framework

Simplest case: indep. obs. $X_1, X_2, \ldots$
2 prob. density or mass functions $f_0$ and $f_i$ known.
When "under control", $X_i$'s iid from $f_0$.
Possibly, at some random time point, $U \geq 1$, $f_i$ governs.

$X_1, X_2, \ldots, X_{V-1}$ iid $f_0$
$X_V, X_{V+1}, \ldots$ iid $f_i$.

We want to determine $V$ with little delay.
Let $Z_i = \log \frac{f_i(X_i)}{f_0(X_i)}$

$S_n = \sum_{i=1}^{n} Z_i, \quad S_0 = 0.$

Under $f_0$, $E(Z_i) < 0$, but under $f_i$, $E(Z_i) > 0$.
This follows because $\log x \leq x - 1$ with equality iff $x = 1$.

So $E_0(Z) = \int \log \frac{f_i(x)}{f_0(x)} f_0(x) \, dx$

$\leq \int \left[ \frac{f_i(x)}{f_0(x)} - 1 \right] f_0(x) \, dx = \int f_i(x) \, dx - \int f_0(x) \, dx$

$= 1 - 1 = 0.$

Similar argument shows $-E_0(Z) < 0$. 

Page 11-12
Cusum Plot.

Plot $S_n$ vs. $n$

$S_n$ vs.

mean path
under control

mean path when problem.

Ex: $X_i$'s iid $N(\mu, \sigma^2)$

$\mu_A$: acceptable $< \mu_A$ reject

$Z = \log \frac{e^{-(X-\mu_A)^2/2\sigma^2}}{e^{-(X-\mu)^2/2\sigma^2}} = \frac{\mu_A - \mu}{\sigma^2} \left[ X - \frac{\mu_A + \mu}{2} \right]$

Plot $S_n = \sum Z_i$; or more simply: $\sum (X_i - \frac{\mu_A + \mu}{2})$.

Alternative plot for normal data: $\sum (X_i - \mu_A)$

When "in control" target = mean. Plot is random walk with mean 0. If mean $\uparrow$, plot shows random drift upward. If mean $\downarrow$, drift downward.
Procedure

Motivation: Suppose you have observed $X_1, X_2, \ldots, X_n$.

$H_0$: $X_i$'s iid $f_0$

Simpler alt. hypothesis $H_1$: For particular $v$, $1 \leq v \leq n$, change to $f_j$ occurs at $v$.

Log likelihood ratio: $H_0 \ vs \ H_0 \quad \sum_{k=v}^{n} Z_k = S_n - S_{v-1}$.

Actual $H_A$ composite: one of the $H_j$'s holds. To test, consider max. of the individuals log LR's:

$$\max_{0 \leq k \leq n} (S_n - S_k) = S_n - \min_{0 \leq k \leq n} S_k.$$  

Procedure: For pre-specified "decision interval" $h$, $h > 0$.

Stopping rule: $\tau = \inf \{ n \geq 1, S_n - \min_{0 \leq j \leq n} S_j > h \}$

"Alarm state" Decide $f_j$ now governs, take appropriate action.

Equivalent approach: Page plot. Set $S'_0 = 0$

$$S'_1 = \max[Z_1, 0], \quad S'_2 = \max[S'_1 + Z_2, 0], \ldots$$

$$S'_i = \max[S'_{i-1} + Z_i, 0].$$

$\tau = \inf \{ n \geq 1, S'_n > h \}$
Fig. 1. Standard control chart. Mean of first thirty results = 0·00.
Mean of second thirty results = 0·20.

Fig. 2. Cumulative sum chart. Mean of first thirty results = 0·00.
Mean of second thirty results = 0·20.

Fig. 3. Control chart and cumulative sum plotted when required.
Plots from Ewan, WP & Kemp, KW (1960)
Biometrika, 47, 363-380.

Fig. 1. Data (Simulated)
\[ X_i \sim N(0, 1) \quad \text{ iid } \quad i = 1, \ldots, 30 \]
\[ \sim N(0.2, 1) \quad \text{ iid } \quad i = 31, \ldots, 60 \]

Fig. 2 Cusum plot
Alt. version \[ \sum X_i \quad \text{ assume 0 is target} \]

Fig. 3 Top panel: Page Plot
but based on \[ \sum (X_i - 10) \]
\[ \text{ ip } \mu_A = 0 \quad \mu_R = 1.20. \]

Cusum & Page plot lead to same decision at same time point.

Cusum plot: history of all past results

Page plot: records only results likely to be relevant to detection of change but has advantage of being bounded below by 0; takes less space.

Change more obvious from Cusum or Page plot than from data plot (Fig. 1).
Siegmund, his $\bar{S}$ is my $S = \Sigma z$; his $b$ is my $h$.

On a Page plot, the local min. at $N_1$, $N_1 + N_2$, etc. would be shifted up to 0 + upper boundary stay at $b$.

In sum, start afresh when test ends on a lower boundary (acceptance test). Process stopped when test ends on upper boundary.
CUSUM + SPRT

cusum: repeated SPRTs w. log-scale boundaries $0 \pm h$

$N_1 = \inf \{ n \geq 1 : S_n \notin (0, h) \}$

If $S_{N_1} \geq h$, $\tau = N_1$

Otherwise, $S_{N_1} = \min_{0 \leq k \leq N_1} \exists S_k \neq 0 \land S_{N_1} \leq 0$

$N_2 = \inf \{ n \geq 1 : S_{N_1+n} - S_{N_1} \notin (0, h) \}$

We have restarted a Wald SPRT at $0$ at $N_1$

$N_2$ = stopping time (starting from $N_1$) for $2^{nd}$ Wald test.

If $S_{N_1+N_2} \geq h$, $\tau = N_1 + N_2$ else $S_{N_1+N_2} \leq S_{N_1}$

$+ S_{N_1+N_2} = \min_{0 \leq k \leq N_1+N_2} \exists S_k \neq 0$

In general, $N_k = \inf \{ n \geq 1 : S_{N_1+...+N_{k-1}+n} - S_{N_1+...+N_{k-1}} \notin (0, h) \}$

$\tau = N_1 + ... + N_{M}$

Where $M = \inf \{ k : S_{N_1+...+N_k} - S_{N_1+...+N_{k-1}} \geq h \}$
Properties

Let $P_v$ denote probability when change from $f_0$ to $f_1$ occurs at observation $v$, $v = 1, 2, \ldots$

$P_0$: prob. when change never occurs

$x_1, x_2, \ldots$, iid $f_0$.

1) $\text{Prob} (\mathcal{C} < \infty) = 1$ under any $P_v$, $v = 1, 2, \ldots, \infty$

2) Optimality. Minimax

Consider all stopping times $N$ w. $E N \geq \delta$

(False alarm rate $= \frac{E_{\mathcal{C}} N}{E_{\mathcal{C}} N} \leq \frac{1}{\delta}$)

Want to minimize

$$\sup_v \text{ess sup } E \left[ (N - v + 1)^+ \mid x_1, \ldots, x_{v-1} \right]$$


ess sup stands for essential supremum.

Note: $E[CN-NV^+ \mid x_1, \ldots, x_{v-1}]$ is fn of $x_i$'s, random

Let $w$ be (possibly vector-valued) r.v. $\circ f(w)$ real valued fn.

$\text{ess sup } f = \inf \{\alpha : P^w (f \not\geq \alpha^+ ) = 0 \}$

A way to apply concept "sup" to random variable.
So want to minimize the longest possible time to detect a change over all possible time points for the change and all possible observations preceding the change.

Page's CUSUM meets the goal.

Moustakides gives proof: nontrivial.

Worst case scenario in Page's scheme occurs when change-point coincides with scheme at an acceptance boundary \((S'_i = 0)\). Obvious why: has farthest to go.

So worst case: \(v = 1\).

For Page CUSUM
\[
\sup_v \text{ess sup } \mathbb{E} \left( (\tau - v + 1)^+ \mid X_1, \ldots, X_{v-1} \right)
\]
\[
= E_j(\tau) \quad \text{ARL under out-of-control regime}
\]

3) MLE of change point \( \tau \) is given by the largest \( n \leq \tau \) for which \( S'_{n-1} = 0 \)

(Assuming we use only obs. up to \( \tau \)).

Start of final Wald SERT