# SEQUENTIAL ESTIMATION OF A RANDOM WALK FIRST-PASSAGE TIME FROM CORRELATED OBSERVATIONS

Marat V. Burnashev

*Institute for Information Transmission Problems,*

*Russian Academy of Sciences, Moscow, Russia*

Aslan Tchamkerten

*Telecom ParisTech, France*

Based on:

1. M.V. Burnashev, A.Tchamkerten,"Sequential Estimation of a Threshold Crossing Moment for a Gaussian Random Walk through Correlated Observations", Probl. of Inform. Trans., 48, no. 2, pp. 65-78, 2012.

2. M.V. Burnashev, A.Tchamkerten,"Estimating a Random Walk First-Passage Time from Noisy or Delayed Observations", IEEE Trans. on Inform. Theory, 58, no. 7, pp. 4230-4243, 2012.

## 1. Problem statement

• Consider the discrete time random process $X = \{X_n\}$

$$X: \quad X_0 = 0, \qquad X_n = sn + \sum_{i=1}^{n} V_i, \qquad n \geq 1,$$

$\{V_i\}$ - indep. $\mathcal{N}(0,1)$–Gaussian r.v.'s, $s > 0$ - given.

• For given threshold level $A > 0$ consider first-passage time

$$\tau_A = \min\{n \geq 0 : X_n \geq A\}.$$

• **We can not** observe process $X$ directly.

• **We observe** another rand. process $Y = \{Y_n\}$.

• Based on observations $Y$ we want to estimate moment $\tau_A$ using stopping time $\eta$ over $Y$ (based on $Y_1, Y_2, \ldots$).

• estimate $\eta$ – Markov stopping time wrt $Y$:
for any $n$ event $\{\eta = n\}$ is defined by observations $Y_1, \ldots, Y_n$ (i.e. $\{\eta = n\} \in \mathcal{F}(Y_1, \ldots, Y_n)$).

• We want to find the value

$$\mathbf{E}|\eta - \tau_A|^p \Longrightarrow \inf_{\eta}, \qquad p > 0$$

Probably, that kind of problem statement (i.e. to estimate a stopping time $\tau$ of $X$ from correlated observations $Y$) appeared first in [Niesen, Tchamkerten, 2009].

Also in [Niesen, Tchamkerten, 2009] a number of possible applications are described (communications, finance, economy, etc.).

**Example**. $X = \{X_t\}$ represents objective price of some shares. For optimal investment we need to know moment $\tau_A$ for some $A > 0$. On stock market we can observe only $Y$ = corrupted version of $X$ (due to human factors, delays, etc.).

Based on observations $Y$ we should estimate the moment $\tau_A$.

• We consider two types of observation process $Y = \{Y_n\}$: **Noisy observations** and **Delayed observations**.

• **Noisy observations**. Observation process $Y$ has form

$$Y: \quad Y_0 = 0, \qquad Y_n = X_n + \varepsilon \sum_{i=1}^{n} W_i, \qquad n \geq 1,$$

$\{W_i\}$ - indep. $\mathcal{N}(0,1)$ – r.v.'s (indep. on $\{V_i\}$), $\varepsilon > 0$ - given.

• For given $A$ and estimate $\eta$ introduce function

$$q(A, p, \eta) = \mathbf{E}|\eta - \tau_A|^p.$$

• We are interested in minimal possible function $q(A, p)$

$$q(A, p) = \inf_{\eta} \mathbf{E}|\eta - \tau_A|^p,$$

inf – over all stopping times $\eta$ wrt $Y$.

⇒ Investigate asymptotics: $s, \varepsilon$ – fixed and $A \to \infty$.

• **Delayed observations**. We are given fixed delay $d = d(A) > 0$ and $Y$ has form

$$Y: \quad Y_0 = Y_1 = \ldots = Y_d = 0; \qquad Y_n = X_{n-d}, \qquad n \geq d+1.$$

Introduce functions

$$q(A, p, d, \eta) = \mathbf{E}|\eta - \tau_A|^p,$$

and

$$q(A, p, d) = \inf_{\eta} q(A, p, d, \eta),$$

inf – over all stopping times $\eta$ wrt $Y$.

⇒ Investigate asymptotics: $s$ – fixed and $d, A \to \infty$.

• That problem – example of traditional estimation problems:

• In observation space $Y$ we are given probab. distributions

$$\{\mathbf{P}_\theta(y),\, \theta \in \Theta\},$$

$\theta$ - unknown parameter.

• Based on observations $\{Y_n\}$ we want to estimate parameter $\theta$.

• But that problem has unpleasant features.

**Standard methodology**

in estimation theory (how solve problems)

• Consider likelihood ratio process (Radon–Nikodim derivative)

$$Z(n,u,\theta_0) = \ln \frac{d\mathbf{P}_{\theta_0+u}}{d\mathbf{P}_{\theta_0}}\left(Y_0^n\right),$$

$\theta_0$ – true parameter value, and $Y_0^n = \{Y_i, 0 \le i \le n\}$.

• Most of estimation problems are reduced to investigation of certain properties of random process $Z(n,u,\theta_0)$.

• In many cases process $Z(n,u,\theta_0)$ has convenient for investigation form (e.g. sum of indep. r.v.'s, stochastic integral, etc.).

• Here parameter $\theta$ = Markov stopping time.

$\Rightarrow$ Process $Z$ has non-convenient form ("Brownian bridge").

# 2. Main results

Introduce value

$$m_p = \mathbf{E}|\xi|^p = \frac{2^{p/2}}{\sqrt{\pi}}\Gamma\left(\frac{p+1}{2}\right), \qquad \xi \sim \mathcal{N}(0,1).$$

## Noisy observations

**Theorem 1**. *For any $p > 0$ the formula holds*

$$q(A,p) = \left[\frac{\varepsilon^2 A}{s^3(1+\varepsilon^2)}\right]^{p/2}(m_p + o(1)), \qquad A \to \infty.$$

Also

**Statement 1**. *For any $p > 0$ the lower bound holds $(A \to \infty)$*

$$q(A,p) \geq \inf_{\eta(Y_0^\infty)} \mathbf{E}|\eta - \tau_A|^p \geq \left[\frac{\varepsilon^2 A}{s^3(1+\varepsilon^2)}\right]^{p/2}(m_p + o(1)).$$

*Note*: inf is taken over all observation process $Y_0^\infty$ up to $n = \infty$.

## Delayed observations

**Theorem 2**. *For any $p > 0$ the formula holds*

$$q(A,d,p) = \left(\frac{d(A)}{s^2}\right)^{p/2}(m_p + o(1)), \qquad A, d \to \infty.$$

## 3. Why do we need $s \neq 0$ ?

**Proposition 1**. *Assume that* $s = 0$. *Then for any* $A > 0$, $\varepsilon > 0$ *and any* $p \geq 1/2$

$$\inf_{\eta\left(Y_0^\infty\right)} \mathbf{E}|\eta - \tau_A|^p = \infty.$$

# 4. Sketch of proofs

**4.1. Theorem 1 - upper bound for** $q(A,p)$.

• Sufficient to choose good estimate $\eta_A$ and evaluate $\mathbf{E}|\eta_A - \tau_A|^p$. Most straightforward way:

• Construct reasonable estimates $\hat{X}_n = \hat{X}_n(Y_0^n), n > 0$ for $\{X_n\}$ and use estimate $\eta_A$

$$\eta_A = \min\{n : \hat{X}_n \geq A\}.$$

• Since all processes – Gaussian, use linear estimates

$$\hat{X}_n = s(1-\alpha)n + \alpha Y_n,$$

where

$$\alpha = \frac{1}{1+\varepsilon^2}.$$

• Then get

**Statement 2**. *The difference $\eta_A - \tau_A$ can be represented as*

$$\eta_A - \tau_A = \sqrt{\frac{\varepsilon^2 A}{s^3(1+\varepsilon^2)}} \left[1 + O\left(\sqrt{\frac{\ln A}{As}}\right)\right] \zeta + \xi_1,$$

*where:*

1) $\zeta \sim \mathcal{N}(0,1)$;

2) *remaining term $\xi_1$ – "small"; in particular,*

$$\mathbf{E}|\xi_1|^p \le C(p) s^{-2p} \left[1 + (sA\ln A)^{p/4}\right], \quad p > 0.$$

From Statement 2 get upper bound (as $A \to \infty$)

$$q(A,p) \le \mathbf{E}|\eta_A - \tau_A|^p = \left[\frac{\varepsilon^2 A}{s^3(1+\varepsilon^2)}\right]^{p/2} (m_p + o(1)).$$

### 4.1. Theorem 1 - lower bound for $\mathbf{E}|\eta - \tau_A|^p$.

• To get lower bound for $\mathbf{E}|\eta - \tau_A|^p$ it is convenient to replace $\mathbf{E}|\eta - \tau_A|^p$ by related inaccuracy $\mathbf{E}|\hat{X}_T - X_T|^p$ when estimating $X_T$ at fixed moment $T$.

• Note that $\mathbf{E}\tau_A \approx A/s$. Introduce fixed time moment $T$

$$T = \frac{A}{s} - ct_0, \qquad t_0 = \sqrt{\frac{2A\ln A}{s^3}}$$

with sufficiently large $c > 0$.

$\Longrightarrow$ with high probability $\tau_A \in (T, T + 2ct_0)$, since

$$\mathbf{P}\left\{\left|\tau_A - \frac{A}{s}\right| \geq ct_0\right\} \lesssim A^{-c^2}.$$

• Replace observation process $Y_n$ by "better" process $Y'_n$

$$Y'_n = \begin{cases} Y_n, & n \leq T, \\ Y_T + X_n - X_T, & n \geq T, \end{cases}$$

i.e. for process $Y'_n, n > T$ additional observation noise $\varepsilon \sum_{i=T+1}^{n} W_i$ disappears after moment $T$.

$\Rightarrow$ Easier to evaluate $\tau_A$ based on $\{Y'_n\}$ than on $\{Y_n\}$.

• If $\tau_A > T \Rightarrow$ difficulty in estimating $\tau_A$ is equivalent to difficulty in estimating $X_T$. In particular, we have

$$\inf_{\eta(Y_0^\infty)} \mathbf{E}|\eta - \tau_A|^p \geq \frac{1}{s^p} \inf_{\mu(Y_T)} \left[\mathbf{E}|\mu - X_T|^p; \tau_A > T\right] (1 + o(1)).$$

After some algebra get Statement 1. $\square$